Fig. 1

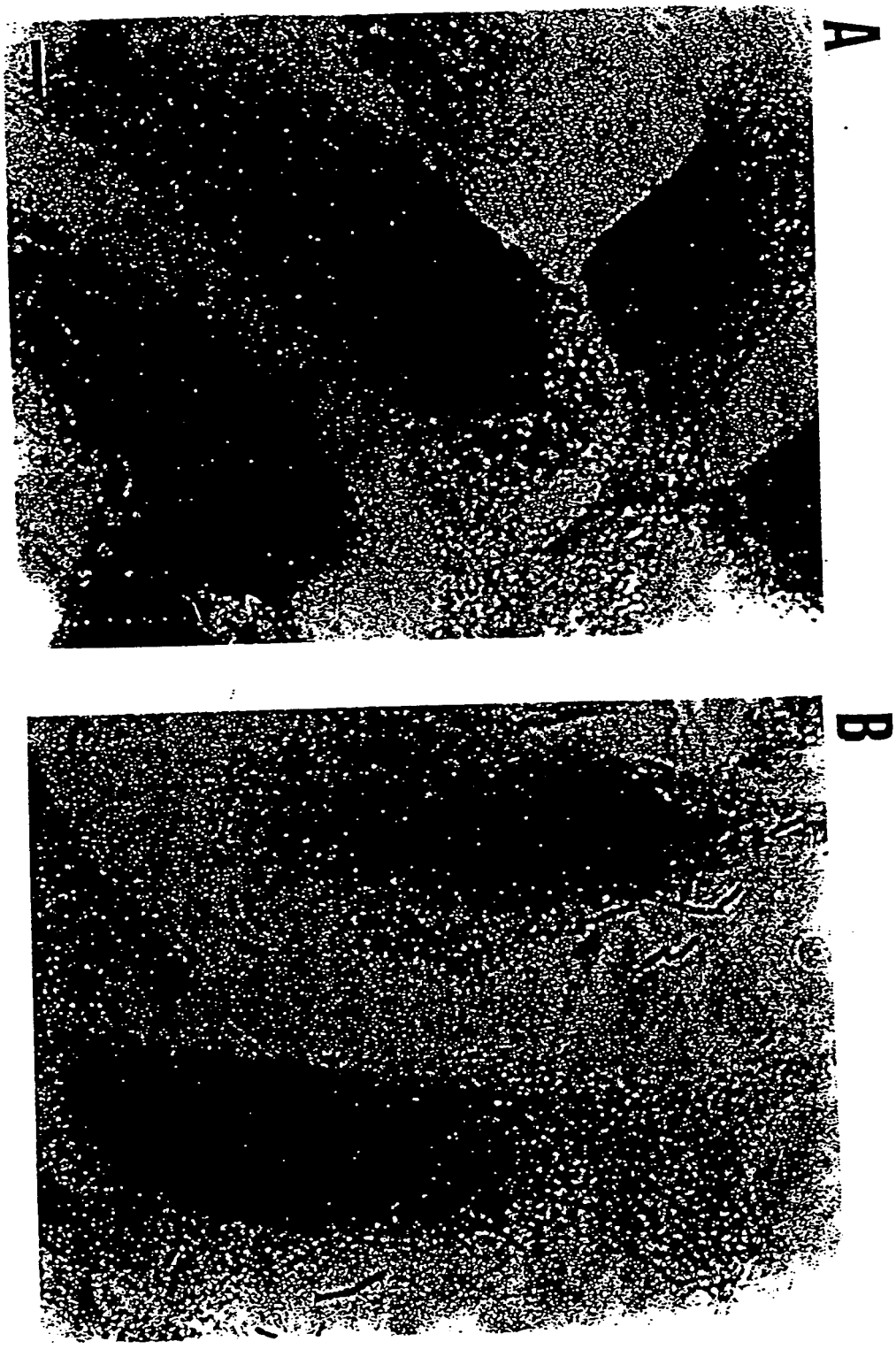


FIGURE 1

FIGURE 2

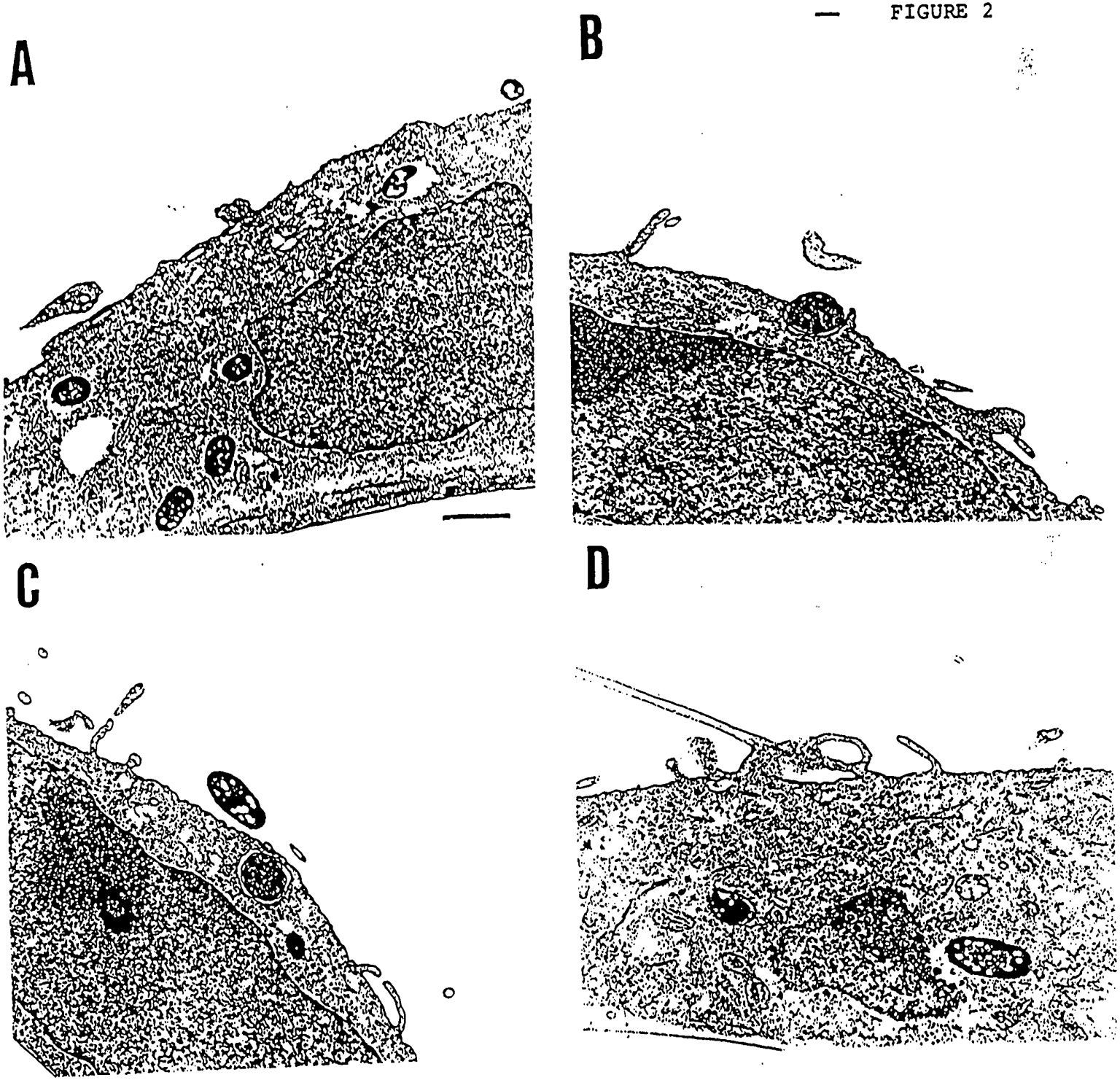


Fig. 3

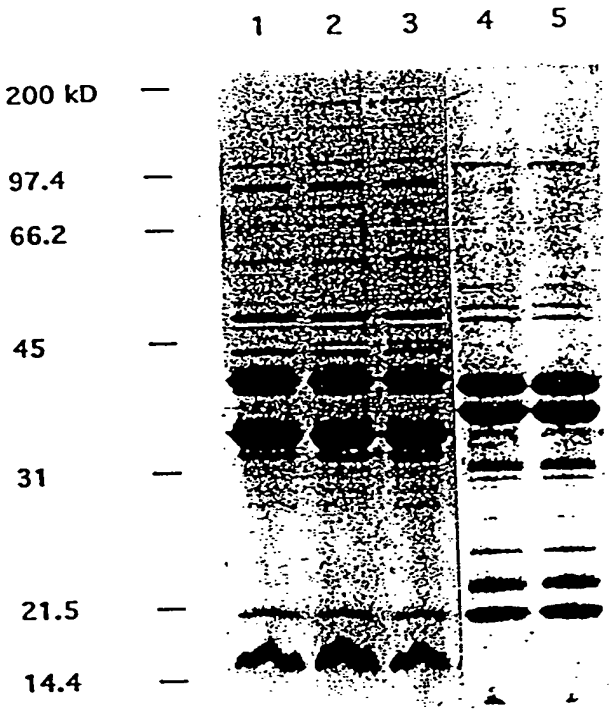


FIGURE 3

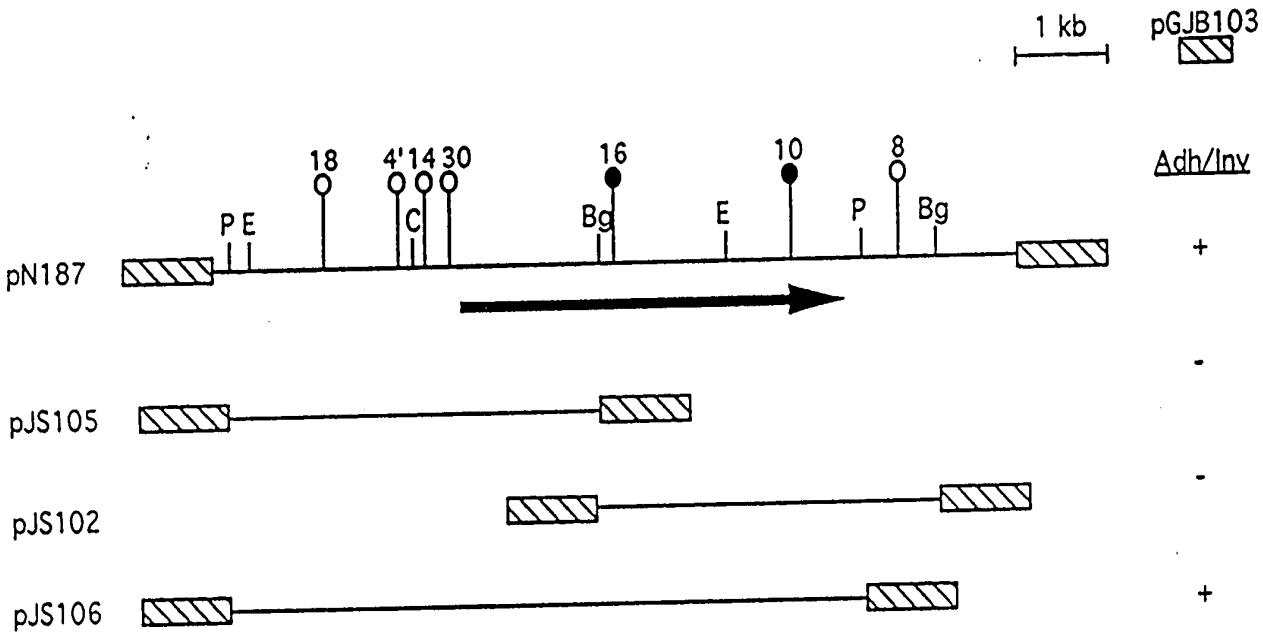


FIGURE 4

Fig. 4

Fig. 5

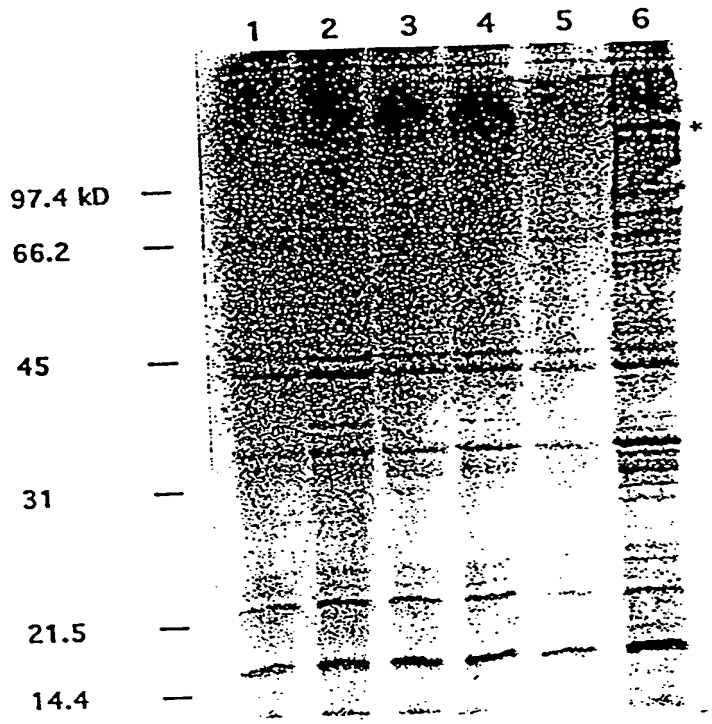


FIGURE 5

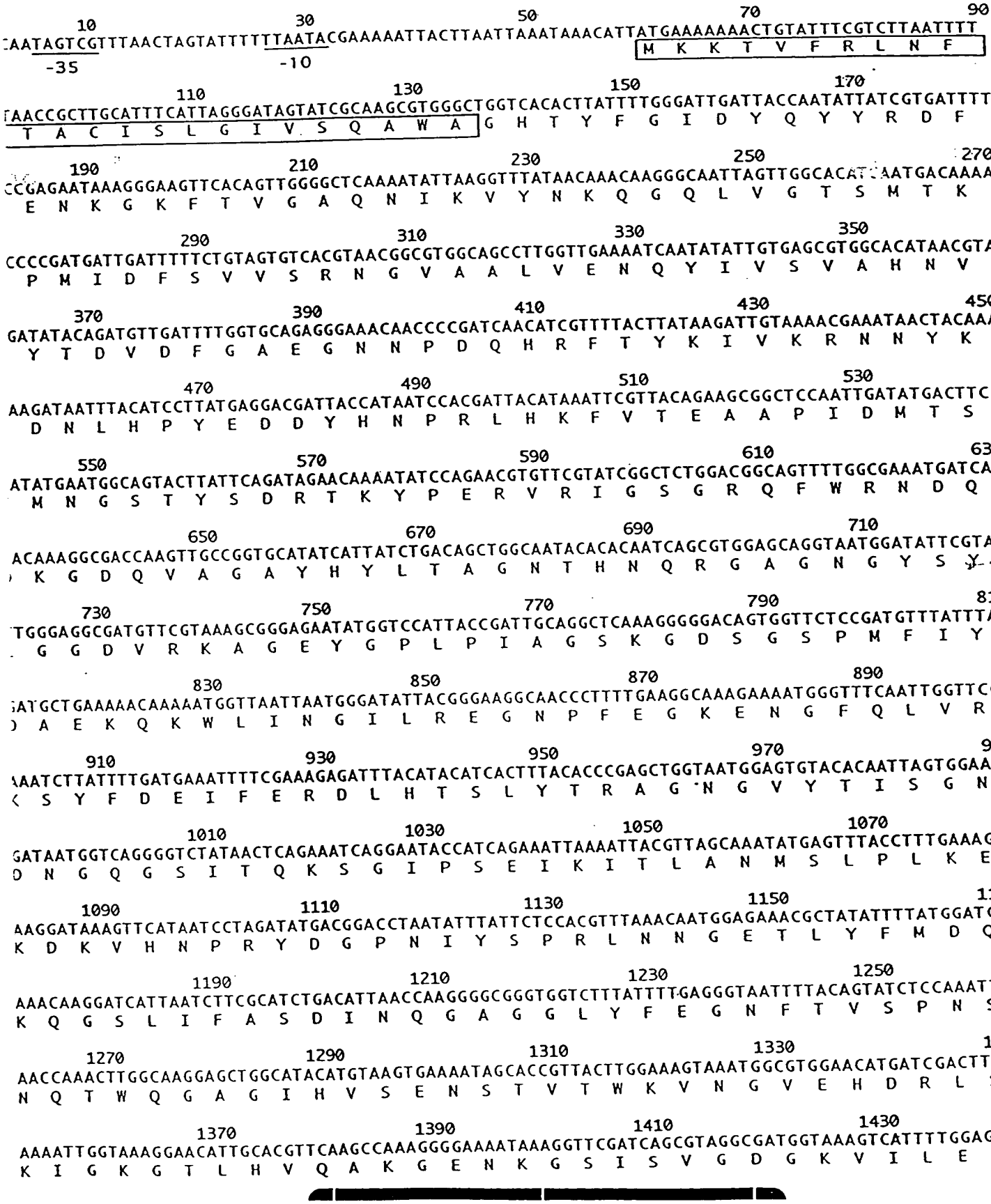


FIGURE 6A

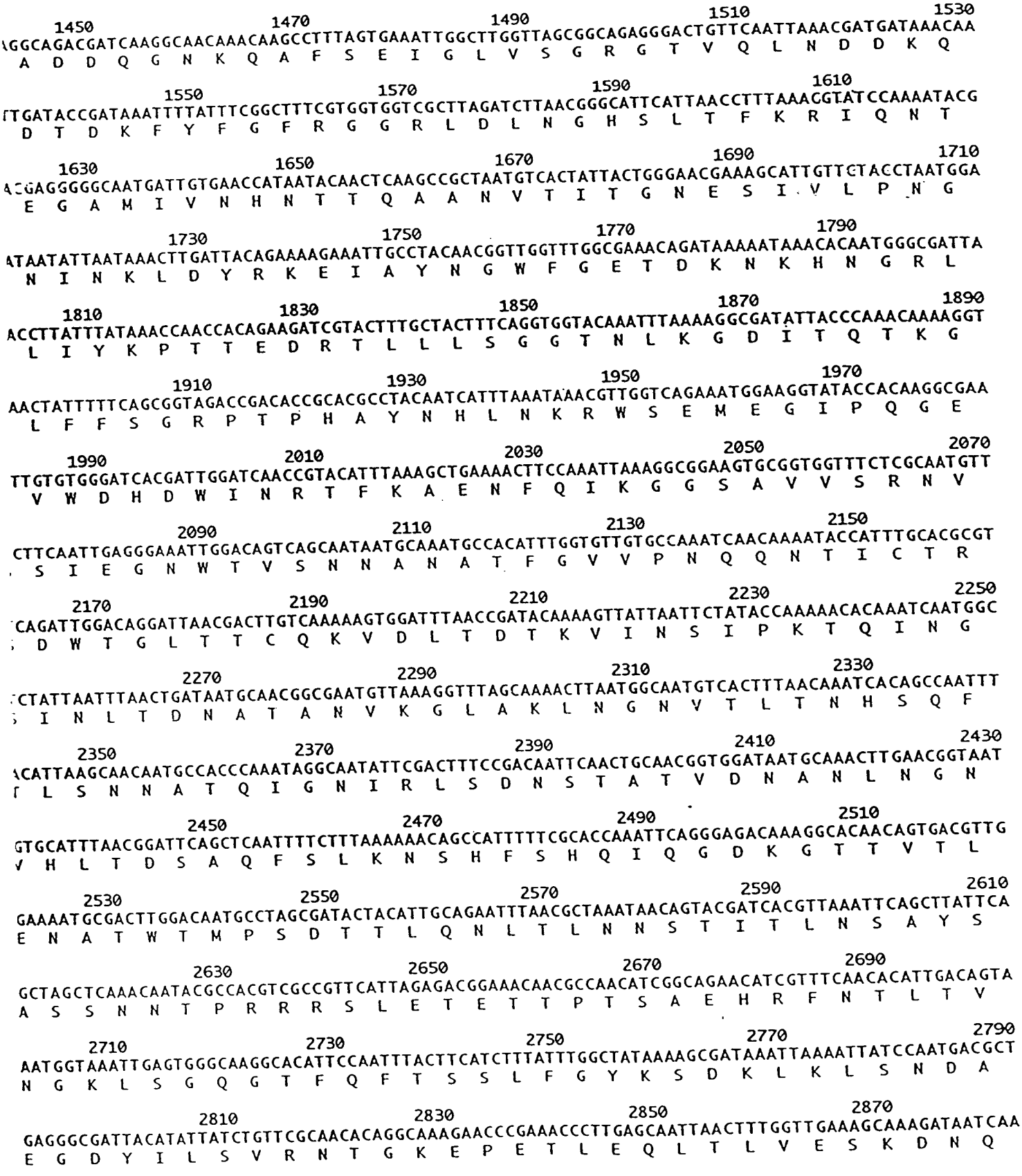

1450 1470 1490 1510 1530
AGGCAGACGATCAAGGCAACAAACAAGCCTTTAGTGAAATTGGCTTGGTTAGCGGCAGAGGGACTGTTCAATTAAACGATGATAAACAA
A D D Q G N K Q A F S E I G L V S G R G T V Q L N D D K Q
1550 1570 1590 1610
TTGATACCGATAAATTTTATTTCGGCTTTCGTGGTGGTCGCTTAGATCTTAACGGGCATTCATTAACCTTTAAAGGTATCCAAAATACG
D T D K F Y F G F R G G R L D L N G H S L T F K R I Q N T
1630 1650 1670 1690 1710
ACGAGGGGGCAATGATTGTGAACCATAATACAACTCAAGCCGCTAATGTCACTATTACTGGGAACGAAAGCATTGTTCTACCTAATGGA
E G A M I V N H N T T Q A A N V T I T G N E S I V L P N G
1730 1750 1770 1790
ATAATATTAATAAACTTGATTACAGAAAAGAAATTGCCTACAACGGTTGGTTTGGCGAAACAGATAAAAATAAACACAATGGGCGATTA
N I N K L D Y R K E I A Y N G W F G E T D K N K H N G R L
1810 1830 1850 1870 1890
ACCTTATTTATAAACCAACCACAGAAGATCGTACTTTGCTACTTTCAGGTGGTACAAATTTAAAAGGCGATATTACCCAAACAAAAGGT
L I Y K P T T E D R T L L L S G G T N L K G D I T Q T K G
1910 1930 1950 1970
AACTATTTTTCAGCGGTAGACCGACACCGCACGCCTACAATCATTTAAATAAACGTTGGTCAGAAATGGAAGGTATACCACAAGGCGAA
L F F S G R P T P H A Y N H L N K R W S E M E G I P Q G E
1990 2010 2030 2050 2070
TTGTGTGGGATCACGATTGGATCAACCGTACATTTAAAGCTGAAAACTTCCAAATTAAAGGCGGAAGTGCGGTGGTTTCTCGCAATGTT
V W D H D W I N R T F K A E N F Q I K G G S A V V S R N V
2090 2110 2130 2150
CTTCAATTGAGGGAAATTGGACAGTCAGCAATAATGCAAATGCCACATTTGGTGTTGTGCCAAATCAACAAAATACCATTTGCACGCGT
S I E G N W T V S N N A N A T F G V V P N Q Q N T I C T R
2170 2190 2210 2230 2250
CAGATTGGACAGGATTAACGACTTGTCAAAAAGTGGATTTAACCGATACAAAAGTTATTAATTCTATACCAAAAACACAAATCAATGGC
D W T G L T T C Q K V D L T D T K V I N S I P K T Q I N G
2270 2290 2310 2330
CTATTAATTTAACTGATAATGCAACGGCGAATGTTAAAGGTTTAGCAAAACTTAATGGCAATGTCACTTTAACAAATCACAGCCAATTT
I N L T D N A T A N V K G L A K L N G N V T L T N H S Q F
2350 2370 2390 2410 2430
ACATTAAGCAACAATGCCACCCAAATAGGCAATATTCGACTTTCCGACAATTCAACTGCAACGGTGGATAATGCAAACTTGAACGGTAAT
L S N N A T Q I G N I R L S D N S T A T V D N A N L N G N
2450 2470 2490 2510
GTGCATTTAACGGATTCAGCTCAATTTTCTTTAAAAAACAGCCATTTTTCGCACCAAATTCAGGGAGACAAAGGCACAACAGTGACGTTG
V H L T D S A Q F S L K N S H F S H Q I Q G D K G T T V T L
2530 2550 2570 2590 2610
GAAAATGCGACTTGGACAATGCCTAGCGATACTACATTGCAGAATTTAACGCTAAATAACAGTACGATCACGTTAAATTCAGCTTATTCA
E N A T W T M P S D T T L Q N L T L N N S T I T L N S A Y S
2630 2650 2670 2690
GCTAGCTCAAACAATACGCCACGTCGCCGTTCATTAGAGACGGAAACAACGCCAACATCGGCAGAACATCGTTTCAACACATTGACAGTA
A S S N N T P R R R S L E T E T T P T S A E H R F N T L T V
2710 2730 2750 2770 2790
AATGGTAAATTGAGTGGGCAAGGCACATTCCAATTTACTTCATCTTTATTTGGCTATAAAAGCGATAAATTAAAATTATCCAATGACGCT
N G K L S G Q G T F Q F T S S L F G Y K S D K L K L S N D A
2810 2830 2850 2870
GAGGGCGATTACATATTATCTGTTCGCAACACAGGCAAAGAACCCGAAACCCTTGAGCAATTAACTTTGGTTGAAAGCAAAGATAATCAA
E G D Y I L S V R N T G K E P E T L E Q L T L V E S K D N Q

FIGURE 6B

```
      2890                2910                2930                2950                2970
CCGTTATCAGATAAGCTCAAATTTACTTTAGAAAATGACCACGTTGATGCAGGTGCATTACGTTATAAATTAGTGAAGAATGATGGCGAA
P  L  S  D  K  L  K  F  T  L  E  N  D  H  V  D  A  G  A  L  R  Y  K  L  V  K  N  D  G  E

            2990                3010                3030                3050
TTCCGCTTGCATAACCCAATAAAAGAGCAGGAATTGCACAATGATTTAGTAAGAGCAGAGCAAGCAGAACGAACATTAGAAGCCAAACAA
F  R  L  H  N  P  I  K  E  Q  E  L  H  N  D  L  V  R  A  E  Q  A  E  R  T  L  E  A  K  Q

      3070                3090                3110                3130                3150
GTTGAACCGACTGCTAAAACACAAACAGGTGAGCCAAAAGTGCGGTCAAGAAGAGCAGCGAGAGCAGCGTTTCCTGATACCCTGCCTGAT
V  E  P  T  A  K  T  Q  T  G  E  P  K  V  R  S  R  R  A  A  R  A  A  F  P  D  T  L  P  D

            3170                3190                3210                3230
CAAAGCCTGTTAAACGCATTAGAAGCCAAACAAGCTGAACTGACTGCTGAAACACAAAAAAGTAAGGCAAAAACAAAAAAAGTGCGGTCA
Q  S  L  L  N  A  L  E  A  K  Q  A  E  L  T  A  E  T  Q  K  S  K  A  K  T  K  K  V  R  S

      3250                3270                3290                3310                3330
AAAAGAGCAGTGTTTTCTGATCCCCTGCTTGATCAAAGCCTGTTCGCATTAGAAGCCGCACTTGAGGTTATTGATGCCCCACAGCAATCG
K  R  A  V  F  S  D  P  L  L  D  Q  S  L  F  A  L  E  A  A  L  E  V  I  D  A  P  Q  Q  S

            3350                3370                3390                3410
GAAAAAGATCGTCTAGCTCAAGAAGAAGCGGAAAAACAACGCAAACAAAAAGACTTGATCAGCCGTTATTCAAATAGTGCGTTATCAGAA
E  K  D  R  L  A  Q  E  E  A  E  K  Q  R  K  Q  K  D  L  I  S  R  Y  S  N  S  A  L  S  E

      3430                3450                3470                3490                3510
TTATCTGCAACAGTAAATAGTATGCTTTCTGTTCAAGATGAATTAGATCGTCTTTTTGTAGATCAAGCACAATCTGCCGTGTGGACAAAT
L  S  A  T  V  N  S  M  L  S  V  Q  D  E  L  D  R  L  F  V  D  Q  A  Q  S  A  V  W  T  N

            3530                3550                3570                3590
ATCGCACAGGATAAAAGACGCTATGATTCTGATGCGTTCCGTGCTTATCAGCAGCAGAAAACGAACTTACGTCAAATTGGGGTGCAAAAA
I  A  Q  D  K  R  R  Y  D  S  D  A  F  R  A  Y  Q  Q  Q  K  T  N  L  R  Q  I  G  V  Q  K

      3610                3630                3650                3670                3690
GCCTTAGCTAATGGACGAATTGGGGCAGTTTTCTCGCATAGCCGTTCAGATAATACCTTTGATGAACAGGTTAAAAATCACGCGACATTA
A  L  A  N  G  R  I  G  A  V  F  S  H  S  R  S  D  N  T  F  D  E  Q  V  K  N  H  A  T  L

            3710                3730                3750                3770
ACGATGATGTCGGGTTTTGCCCAATATCAATGGGGCGATTTACAATTTGGTGTAAACGTGGGAACGGGAATCAGTGCGAGTAAAATGGCT
T  M  M  S  G  F  A  Q  Y  Q  W  G  D  L  Q  F  G  V  N  V  G  T  G  I  S  A  S  K  M  A

      3790                3810                3830                3850                3870
GAAGAACAAAGCCGAAAAATTCATCGAAAAGCGATAAATTATGGCGTGAATGCAAGTTATCAGTTCCGTTTAGGGCAATTGGGCATTCAG
E  E  Q  S  R  K  I  H  R  K  A  I  N  Y  G  V  N  A  S  Y  Q  F  R  L  G  Q  L  G  I  Q

            3890                3910                3930                3950
CCTTATTTTGGAGTTAATCGCTATTTTATTGAACGTGAAAATTATCAATCTGAGGAAGTGAGAGTGAAAACGCCTAGCCTTGCATTTAAT
P  Y  F  G  V  N  R  Y  Y  F  I  E  R  E  N  Y  Q  S  E  E  V  R  V  K  T  P  S  L  A  F  N

      3970                3990                4010                4030                4050
CGCTATAATGCTGGCATTCGAGTTGATTATACATTTACTCCGACAGATAATATCAGCGTTAAGCCTTATTTCTTCGTCAATTATGTTGAT
R  Y  N  A  G  I  R  V  D  Y  T  F  T  P  T  D  N  I  S  V  K  P  Y  F  F  V  N  Y  V  D

            4070                4090                4110                4130
GTTTCAAACGCTAACGTACAAACCACGGTAAATCTCACGGTGTTGCAACAACCATTTGGACGTTATTGGCAAAAAGAAGTGGGATTAAAG
V  S  N  A  N  V  Q  T  T  V  N  L  T  V  L  Q  Q  P  F  G  R  Y  W  Q  K  E  V  G  L  K

      4150                4170                4190                4210                4230
GCAGAAATTTTACATTTCCAAATTTCCGCTTTTATCTCAAAATCTCAAGGTTCACAACTCGGCAAACAGCAAAATGTGGGCGTGAAATTG
A  E  I  L  H  F  Q  I  S  A  F  I  S  K  S  Q  G  S  Q  L  G  K  Q  Q  N  V  G  V  K  L

            4250                4270                4290                4310
GGCTATCGTTGGTAAAAATCAACATAATTTTATCGTTTATTGATAAACAAGGTGGGTCAGATCAGATCCCACCTTTTTTATTCCAATAAT
G  Y  R  W  *
```

FIGURE 6C

Fig. 7

```
                1                                                   50
Hap             MKKTVFRLNF LTACISLGIV SQAWAGHTYF GIDYQYYRDF AENKGKFTVG
HK368IGA        MLNKKFKLNF IALTVAYALT PYTEAALVRD DVDYQIFRDF AENKGKFSVG
HK393IGA        MLNKKFKLNF IALTVAYALT PYTEAALVRD DVDYQIFRDF AENKGKFSVG
HK715IGA        MLNKKFKLNF IALTVAYALT PYTEAALVRD DVDYQIFRDF AENKGRFSVG
HK61IGA         MLNKKFKLNF IALTVAYALT PYTEAALVRD DVDYQIFRDF AENKGKFSVG
Consensus       M----F-LNF ---------- ----A----- --DYQ--RDF AENKG-F-VG

                51                                                  100
Hap             AQNIKVYNKQ GQLVGTSMTK A.PMIDFSVV SRNG.VAALV ENQYIVSVAH
HK368IGA        ATNVLVKDKN NKDLGTALPN GIPMIDFSVV DVDKRIATLI NPQYVVGVKH
HK393IG         ATNVEVRDKN NRPLGNVLPN GIPMIDFSVV DVDKRIATLV NPQYVVGVKH
HK715IGA        ATNVEVRDKN NHSLGNVLPN GIPMIDFSVV DVDKRIATLI NPQYVVGVKH
HK61IGA         ATNVEVRDKK NQSLGSALPN GIPMIDFSVV DVDKRIATLV NPQYVVGVKH
Consensus       A-N--V--K- ----G----- --PMIDFSVV ------A-L- --QY-V-V-H

                101                                                 150
Hap             ......NVGY TDVDFGAEGN NPDQHR.... ..FTYKIVKR NNY.......
HK368IGA        VSNGVSELHF GNLNGNMNNG NAKAHRDVSS EENRYFSVEK NEYPTKLNGK
HK393IGA        VSNGVSELHF GNLNGNMNNG NAKAHRDVSS EENRYYTVEK NEYPTKLNGK
HK715IGA        VSNGVSELHF GNLNGNMNNG NDKSHRDVSS EENRYFSVEK NEYPTKLNGK
HK61IGA         VSNGVSELHF GNLNGNMNNG NAKSHRDVSS EENRYYTVEK NNFPTENVTS
Consensus       ---------- ---------- N---HR---- ----Y--V-- N---------

                151                                                 200
Hap             ....KKDNLH PYEDDYHNPR LHKFVTEAAP IDM.TSNMNG STYSDRTKYP
HK368IGA        TVTTEDQ.TQ KRREDYYMPR LDKFVTEVAP IEASTASSDA GTYNDQNKYP
HK393IGA        AVTTEDQ.AQ KRREDYYMPR LDKFVTEVAP IEASTDSSTA GTYNNKDKYP
HK715IGA        AVTTEDQ.TQ KRREDYYMPR LDKFVTEVAP IEASTASSDA GTYNDQNKYP
HK61IGA         FTTKEEQDAQ KRREDYYMPR LDKFVTEVAP IEASTANNNK GEYNNSDKYP
Consensus       ---------- ----DY--PR L-KFVTE-AP I---T----- --Y----KYP

                201                                                 250
Hap             ERVRIGSGRQ F......... .......... .....WRNDQ DKGDQVAGAY
HK368IGA        AFVRLGSGSQ FIYKKGDNYS LIL......N NH....EVGG NNLKLVGDAY
HK393IGA        YFVRLGSGTQ FIYENGTRYE LWL......G KEGQKSDAGG YNLKLVGNAY
HK715IGA        AFVRLGSGSQ FIYKKGDNYS LIL......N NH....EVGG NNLKLVGDAY
HK61IGA         AFVRLGSGSQ FIYKKGSRYQ LILTEKDKQG NLLRNWDVGG DNLELVGNAY
Consensus       --VR-GSG-Q F--------- ---------- ---------- -----V--AY
```

FIGURE 7A

```
                                                                    300
            251
Hap         HYLTAGNTHN QRGAGNGYSY LGG......D VRKAGEYGPL PIAGSKGDSG
HK368IGA    TYGIAGTPYK VNHENNGLIG FGNSKEEHSD PKGILSQDPL TNYAVLGDSG
HK393IGA    TYGIAGTPYE VNHENDGLIG FGNSNNEYIN PKEILSKKPL TNYAVLGDSG
HK715IGA    TYGIAGTPYK VNHENNGLIG FGNSKEEHSD PKGILSQDPL TNYAVLGDSG
HK61IGA     TYGIAGTPYK VNHENNGLIG FGNSKEEHSD PKGILSQDPL TNYAVLGDSG
Consensus   -Y--AG---- ------G--- -G-------- --------PL -------GDSG
                                                                    *
                                                                    350
            301
Hap         SPMFIYDAEK QKWLINGILR EGNPFEGKEN GFQLVRKSYF D.EIFERDLH
HK368IGA    SPLFVYDREK GKWLFLGSYD FWAGYN.... .....KKSWQ EWNIYKSQFT
HK393IGA    SPLFVYDREK GKWLFLGSYD YWAGYN.... .....KKSWQ EWNIYKPEFA
HK715IGA    SPLFVYDREK GKWLFLGSYD FWAGYN.... .....KKSWQ EWNIYKPEFA
HK61IGA     SPLFVYDREK GKWLFLGSYD FWAGYN.... .....KKSWQ EWNIYKHEFA
Consensus   SP-F-YD-EK -KWL--G--- ---------- ------KS-- ---I------
                                                                    400
            351
Hap         TSLYTRAGNG VYTISGNDNG QGSITQKSGI PSEIKITLAN MSLPLKEKDK
HK368IGA    KDVLNKDSAG SLIGSKIDYS WSSNGKTSTI TGGEK....S LNVDLAD...
HK393IGA    EKIYEQYSAG SLIGSKIDYS WSSNGKTSTI TGGEK....S LNVDLAD...
HK715IGA    KTVLDKDTAG SLTGSNTQYN WNPTGKTSVI SNGSE....S LNVDLFD...
HK61IGA     EKIYQQYSAG SLTGSNTQYT WQATGSTSTI TGGGE....P LSVDLTD...
Consensus   ---------G ----S----- -------S-I ---------- ----L-----
                                                                    450
            401
Hap         VHNPRYDGPN IYSPRLNNGE TLYFMDQKQG SLIFASDINQ GAGGLYFEGN
HK368IGA    ......GKD. ....KPNHGK SVIFEG..SG TLTLNNNIDQ GAGGLFFEGD
HK393IGA    ......GKD. ....KPNHGK SVIFEG..SG TLTLNNNIDQ GAGGLFFEGD
HK715IGA    ......SSQD TDSKKNNHGK SVTLRG..SG TLTLNNNIDQ GAGGLFFEGD
HK61IGA     ......GKD. ....KPNHGK SITLKG..SG TLTLNNHIDQ GAGGLFFEGD
Consensus   ---------- ------N-G- ---------G -L-----I-Q GAGGL-FEG-
                                                                    500
            451
Hap         FTVSPNSNQ. TWQGAGIHVS ENSTVIWKVN GVEHDRLSKI GKGTLHVQAK
HK368IGA    YEVKGTSDNT TWKGAGVSVA EGKTVIWKVH NPQYDRLAKI GKGTLIVEGT
Hk393IGA    YEVKGTSDNT TWKGAGVSVA EGKTVIWKVH NPQYDRLAKI GKGTLIVEGT
HK715IGA    YEVKGTSDST TWKGAGVSVA DGKTVIWKVH NPKSDRLAKI GKGTLIVEGK
HK61IGA     YEVKGTSDST TWKGAGVSVA DGKTVIWKVH NPKYDRLAKI GKGTLVVEGK
Consensus   --V---S--- TW-GAG--V- ---TVIWKV- ----DRL-KI GKGTL-V---
```

FIGURE 7B

```
                                                                    550
           501
Hap        GENKGSISVG DGKVILEQQA DDQGNKQAFS EIGLVSGRGT VQLNDDKQFD
HK368IGA   GDNKGSLKVG DGTVILKQQT NGSGQ.HAFA SVGIVSGRST LVLNDDKQVD
HK393IGA   GDNKGSLKVG DGTVILKQQT NGSGQ.HAFA SVGIVSGRST LVLNDDKQVD
HK715IGA   GENKGSLKVG DGTVILKQQA DANNKVKAFS QVGIVSGRST VVLNDDKQVD
HK61IGA    GKNEGLLKVG DGTVILKQKA DANNKVQAFS QVGIVSGRST LVLNDDKQVD
Consensus  G-N-G---VG DG-VIL-Q-- -------AF- --G-VSGR-T---LNDDKQ-D

                                                                    600
           551
Hap        TDKFYFGFRG GRLDLNGHSL TFKRIQNIDE GAMIVNHNIT QAANVTITGN
HK368IGA   PNSIYFGFRG GRLDLNGNSL TFDHIRNIDD GARLVNHNMT NASNITITGE
HK393IGA   PNSIYFGFRG GRLDLNGNSL TFDHIRNIDE GARLVNHSTS KHSTVTITGD
HK715IGA   PNSIYFGFRG GRLDANGNNL TFEHIRNIDD GARLVNHNTS KTSTVTITGE
HK61IGA    PNSIYFGFRG GRLDLNGNSL TFDHIRNIDD GARVVNHNMT NTSNITITGE
Consensus  ----YFGFRG GRLD-NG--L TF--I-N-D- GA--VNH--- -----TITG-

                                                                    650
           601
Hap        ESIVLPNG.. .......... .......... .......... ..........
HK368IGA   SLITDPNTIT PYNIDAPDED NPYAFRRIKD GGQLYLNLEN YTYYALRKGA
HK393IGA   NLITDPNNVS IYYVKPLEDD NPYAIRQIKY GYQLYFNEEN RTYYALKKDA
HK715IGA   SLITDPNTIT PYNIDAPDED NPYAFRRIKD GGQLYLNLEN YTYYALRKGA
HK61IGA    SLITNPNTIT SYNIEAQDDD HPLRIRSIPY R.QLYFNQDN RSYYTLKKGA
Consensus  --I--PN--- ---------- ---------- ---------- ----------

                                                                    700
           651
Hap        .......... .......... .........N NINKLDYRKE IAYNGWFGET
HK368IGA   STRSELPKNS GESNENWLYM GKTSDEAKRN VMNHINNERM NGFNGYFGEE
HK393IGA   SIRSEFPQNR GESNNSWLYM GTEKADAQKN AMNHINNERM NGFNGYFGEE
HK715IGA   STRSELPKNS GESNENWLYM GKTSDEAKRN VMNHINNERM NGFNGYFGEE
HK61IGA    STRSELPQNS GESNENWLYM GRTSDEAKRN VMNHINNERM NGFNGYFGEE
Consensus  ---------- ---------- ---------N --N------- ---NG-FGE-

                                                                    750
           701
Hap        D.KNKHNGRL NLIYKPTTED RTLLLSGGTN LKGDITQTKG KLFFSGRPTP
HK368IGA   EGK..NNGNL NVTFKGKSEQ NRFLLTGGTN LNGDLTVEKG TLFLSGRPTP
HK393IGA   EGK..NNGNL NVTFKGKSEQ NRFLLTGGTN LNGDLNVQQG TLFLSGRPTP
HK715IGA   EGK..NNGNL NVTFKGKSEQ NRFLLTGGTN LNGDLKVEKG TLFLSGRPTP
HK61IGA    ETKATQNGKL NVTFNGKSDQ NRFLLTGGTN LNGDLNVEKG TLFLSGRPTP
Consensus  --K---NG-L N---------- ---LL-GGTN L-GD-----G -LF-SGRPTP
```

FIGURE 7C

```
                                                                    800
           751
Hap        HAYNHLNKRW SEMEG..IPQ GEIVWDHDWI NRTFKAENFQ IKGGSAVVS.
HK368IGA   HARDIAGISS TKKDPHFAEN NEVVVEDDWI NRNFKATTMN VTGNASLYSG
HK393IGA   HARDIAGISS TKKDSHFSEN NEVVVEDDWI NRNFKATNIN VTNNATLYSG
HK715IGA   HARDIAGISS TKKDQHFAEN NEVVVEDDWI NRNFKATNIN VTNNATLYSG
HK61IGA    HARDIAGISS TKKDPHFTEN NEVVVEDDWI NRNFKATTMN VTGNASLYSG
Consensus  HA-------- ---------- -E-V---DWI NR-FKA---- --------S-

                                                                    850
           801
Hap        RNVSSIEGNW TVSNNANATF GVVPNQQNTI CTRSDWIGLT TCQKVDLTDT
HK368IGA   RNVANITSNI TASNKAQVHI GY..KTGDTV CVRSDYTGYV TCTTDKLSD.
HK393IGA   RNVESITSNI TASNNAKVHI GY..KAGDTV CVRSDYTGYV TCTTDKLSD.
HK715IGA   RNVANITSNI TASDNAKVHI GY..KAGDTV CVRSDYTGYV TCTTDKLSD.
HK61IGA    RNVANITSNI TASNNAQVHI GY..KTGDTV CVRSDYTGYV TCHNSNLSE.
Consensus  RNV--I--N- T-S--A---- G-------T- C-RSD-TG-- TC----L---
                                             *           *

                                                                    900
           851
Hap        KVINSIPKTQ INGSINLTDN ATANVKGLAK LNGNVTLTNH SQFTLSNNAT
HK368IGA   KALNSFNPTN LRGNVNLTES A......... .......... ..........
HK393IGA   KALNSFNPTN LRGNVNLTES A......... .......... ..........
HK715IGA   KALNSFNATN VSGNVNLSGN A......... .......... ..........
HK61IGA    KALNSFNPTN LRGNVNLTEN A......... .......... ..........
Consensus  K--NS---T- --G--NL--- A--------- ---------- ----------

                                                                    950
           901
Hap        QIGNIRLSDN STATVDNANL NGNVHLTDSA QFSLKNSHFS HQIQGDKGTT
HK368IGA   .......... .NFVLGKANL FGTIQSRGNS QVRLT..... ..........
HK393IGA   .......... .NFVLGKANL FGTIQSRGNS QVRLT..... ..........
HK715IGA   .......... .NFVLGKANL FGTISGTGNS QVRLT..... ..........
HK61IGA    .......... .SFTLGKANL FGTIQSIGTS QVNLK..... ..........
Consensus  ---------- -------ANL -G-------- Q--L------ ----------

                                                                   1000
           951
Hap        VTLENATWIM PSDTTLQNLT LNNSTITLNS AYSASSNNTP RRRSLETETT
HK368IGA   ...ENSHWHL TGNSDVHQLD LANGHIHLNS ADNSNNVTK. ..........
HK393IGA   ...ENSHWHL TGNSDVHQLD LANGHIHLNS ADNSNNVTK. ..........
HK715IGA   ...ENSHWHL TGDSNVNQLN LDKGHIHLNA QNDANKVTT. ..........
HK61IGA    ...ENSHWHL TGNSNVNQLN LTNGHIHLNA QNDANKVTT. ..........
Consensus  ---EN--W-- --------L- L----I-LN- ---------- ----------
```

FIGURE 7D

```
                1001                                                1050
Hap             PTSAEHRFNT LTVNGKLSGQ GTFQFTSSLF GYKSDKLKLS NDAEGDYILS
HK368IGA        .......YNT LTVNS.LSGN GSFYYLTDLS NKQGDKVVVT KSATGNFTLQ
HK393IGA        .......YNT LTVNS.LSGN GSFYYLTDLS NKQGDKVVVT KSATGNFTLQ
HK715IGA        .......YNT LTVNS.LSGN GSFYYLTDLS NKQGDKVVVT KSATGNFTLQ
HK61IGA         .......YNT LTVNS.LSGN GSFYYWVDFT NNKSNKVVVN KSATGNFTLQ
Consensus       --------NT LTVN--LSG- G-F------- -----K---- --A-G---L-

                1051                                                1100
Hap             VRNTGKEPET LEQLTLVESK DNQPLSDKLK FTLENDHVDA GALRYKLVKN
HK368IGA        VADKTGEPNH .NELTLFDAS KAQR..DHLN VSLVGNTVDL GAWKYKLRNV
HK393IGA        VADKTGEPNH .NELTLFDAS KAQR..DHLN VSLVGNTVDL GAWKYKLRNV
HK715IGA        VADKTGEPTK .NELTLFDAS NATR..NNLN VSLVGNTVDL GAWKYKLRNV
HK61IGA         VADKTGEPNH .NELTLFDAS NATR..NNLE VTLANGSVDR GAWKYKLRNV
Consensus       V-----EP-- ---LTL---- --------L- --L----VD- GA--YKL---

                1101                                                1150
Hap             DGEFRLHNPI KEQELHNDLV .......... .......... ..........
HK368IGA        NGRYDLYNP. .EVEKRNQTV DTTNITTPNN IQADVPSVPS NNEEIARVDE
HK393IGA        NGRYDLYNP. .EVEKRNQTV DTTNITTPNN IQADVPSVPS NNEEIARVDE
HK715IGA        NGRYDLYNP. .EVEKRNQTV DTTNITTPNN IQADVPSVPS NNEEIARV.E
HK61IGA         NGRYDLYNP. .EVEKRNQTV DTTNITTPND IQADAPSAQS NNEEIARV.E
Consensus       -G---L-NP- -E-E--N--V ---------- ---------- ----------

                1151                                                1200
Hap             .......... .......... .......... .......... ..........
HK368IGA        APVPPPAPAT .......... .......... .......... ..........
HK393IGA        APVPPPAPAT .......... .......... .......... ..........
HK715IGA        TPVPPPAPAT .......... .......... .......... ..........
HK61IGA         TPVPPPAPAT ESAIASEQPE TRPAETAQPA MEETNTANST ETAPKSDTAT
Consensus       ---------- ---------- ---------- ---------- ----------

                1201                                                1250
Hap             .......... .......... RAEQAERTLE AKQVEPT... ..........
HK368IGA        .......... PSETTETVAE NSKQESKTVE KNEQDATETT AQNREVAKEA
HK393IGA        .......... PSETTETVAE NSKQESKTVE KNEQDATETT AQNREVAKEA
HK715IGA        .......... PSETTETVAE NSKQESKTVE KNEQDATETT AQNGEVAEEA
HK61IGA         QTENPNSESV PSETTEKVAE NPPQENETVA KNEQEATEPT PQNGEVAKED
Consensus       ---------- ---------- ---Q---T-- ------T--- ----------
```

FIGURE 7E

```
                                                                    1300
            1251
Hap         .....AKTQT GE..... ..... ..... .....
HK368IGA    KSNVKANTQT NEVAQSGSET KETQTTETK. ..... ....ETATVE
HK393IGA    KSNVKANTQT NEVAQSGSET KETQTTETK. ..... ....ETATVE
HK715IGA    KPSVKANTQT NEVAQSGSET EETQTTEIK. ..... ....ETAKVE
HK61IGA     QPTVEANTQT NEATQSEGKT EETQTAETKS EPTESVTVSE NQPEKTVSQS
Consensus   -----A-TQT -E-------- ---------- -- -------- ----------

                                                                    1350
            1301
Hap         ..... ..... ..... ..... .....
HK368IGA    KEEK.. ..... ..... ..... .....
HK393IGA    KEEK.. ..... ..... ..... .....
HK715IGA    KEEKAKVEKE EKAKVEKDEI QEAPQMASET SPKQAKPAPK EVSTDTKVEE
HK61IGA     TEDKVVVEKE EKAKVETEET QKAPQVTSKE PPKQAEPAPE EVPTDTNAEE
Consensus   ---------- ---------- ---------- ---------- ----------

                                                                    1400
            1351
Hap         ..... ..... ..... ..... .....
HK368IGA    ..... ..... ..... ..... .....
HK393IGA    ..... ..... ..... ..... .....
HK715IGA    TQVQAQPQTQ STTVAAAEAT SPNSKPAEET .QPSEKTNAE PVTPVVSKNQ
HK61IGA     A..QALQQTQ PTTVAAAETT SPNSKPAEET QQPSEKTNAE PVTPVVS...
Consensus   ---------- ---------- ---------- ---------- ----------

                                                                    1450
            1401
Hap         ..... ..... .....PKVRS RRAARAAFPD TLP.....
HK368IGA    ..... ....AKVETE KTQEVPKVTS QVSPKQEQSE T.....
HK393IGA    ..... ....AKVETE KTQEVPKVTS QVSPKQEQSE T.....
HK715IGA    TENTIDQPTE REKTAKVETE KTQEPPQVAS QASPKQEQSE T.....
HK61IGA     .ENTATQPTE TEETAKVEKE KTQEVPQVAS QESPKQEQPA AKPQAQTKPQ
Consensus   ---------- ---------- -----P-V-S ---------- ----------

                                                                    1500
            1451
Hap         ..... ..... ..... ..... .....V
HK368IGA    ..... ..... ..... ..... .....V
HK393IGA    ..... ..... ..... ..... .....V
HK715IGA    ..... ..... ..... ..... .....V
HK61IGA     AEPARENVLT TKNVGEPQPQ AQPQTQSTAV PTTGETAANS KPAAKPQAQA
Consensus   ---------- ---------- ---------- ---------- ----------
```

FIGURE 7F

```
                                                                    1550
            1501
Hap         .........D QSLLNALEA. .....KQAEL TAETQKSKAK TKK.......
HK368IGA    QPQAEPAREN DPTVNIKEP. ....QSQTNT TADTEQPAKE TSSNVE....
HK393IGA    QPQAEPAREN DPTVNIKEP. ....QSQTNT TADTEQPAKE TSSNVE....
HK715IGA    QPQAVLESEN VPTVNNAEEV QAQLQTQTSA TVSTKQPAPE NSINTG....
HK61IGA     KPQTEPAREN VSTVNTKEP. ....QSQTSA TVSTEQPAKE TSSNVEQPAP
Consensus   ---------- ----N--E-- ------Q--- T--T------ ----------

                                                                    1600
            1551
Hap         .......... .......... .........V RSKRAVFSDP LLDQSL....
HK368IGA    .......... .......... .......QPVT ESTTVNTGNS VVEN......
HK393IGA    .......... .......... .......QPVT ESTTVNTGNS VVEN......
HK715IGA    .......SAT AITETAEKSD KPQTETAAST EDASQHKANT VADNSVANNS
HK61IGA     ENSINTGSAT TMTETAEKSD KPQMET..VT ENDRQPEANT VADNSVANNS
Consensus   ---------- ---------- ---------- ---------- ----------

                                                                    1650
            1601
Hap         .......... .........F ALEAALEVID APQQSEKDRL AQEEAEKQRK
HK368IGA    .......... .......... PENTTPATTQ PTVNSESSN. .KPK.NRHRR
HK393IGA    .......... .......... PENTTPATTQ PTVNSESSN. .KPK.NRHRR
HK715IGA    ESSEPKSRRR RSISQPQETS AEETTAASTD ETTIADNSKR SKPN.RRSRR
HK61IGA     ESSESKSRRR RSVSQPKETS AEETTVASTQ ETTVDNSVST PKPRSRRTRR
Consensus   ---------- ---------- ---------- ---------- --------R-

                                                                    1700
            1651
Hap         .......... .......... .......... .....QKDLI SRYSNSALSE
HK368IGA    SVRSVPHNVE PATTSSND.. .......... RSTVALCDLT STNTNAVLSD
HK393IGA    SVRSVPHNVE PATTSSND.. .......... RSTVALCDLT STNTNAVLSD
HK715IGA    SVRS.....E PTVTNGSD.. .......... RSTVALRDLT STNTNAVISD
HK61IGA     SVQTNSYEPV ELPTENAENA ENVQSGNNVA NSQPALRNLT SKNTNAVLSN
Consensus   ---------- ---------- ---------- --------L- S---N---S-

                                                                    1750
            1701
Hap         LSA.....TV NSMLSVQDEL DRL.FVDQAQ SAVWTNIAQD KRRYDSDAFR
HK368IGA    ARAKAQFVAL NVGKAVSQHI SQLEMNNEGQ YNVWVSNTSM NKNYSSSQYR
HK393IGA    ARAKAQFVAL NVGKAVSQHI SQLEMNNEGQ YNVWVSNTSM NKNYSSSQYR
HK715IGA    AMAKAQFVAL NVGKAVSQHI SQLEMNNEGQ YNVWVSNTSM NENYSSSQYR
HK61IGA     AMAKAQFVAL NVGKAVSQHI SQLEMNNEGQ YNVWISNTSM NKNYSSEQYR
Consensus   --A------- N----V---- --L------Q --VW------ ---Y-S---R
```

FIGURE 7G

```
                                                                    1800
            1751
Hap         AYQQQKINLR QIGVQKALAN GRIGAVFSHS RSDNTFDEQV KNHATLTMMS
HK368IGA    RFSSKSTQTQ LGWDQTISNN VQLGGVFTYV RNSNNFDKAT SKN.TLAQVN
HK393IGA    RFSSKSTQTQ LGWDQTISNN VQLGGVFTYV RNSNNFDKAT SKN.TLAQVN
HK715IGA    RFSSKSTQTQ LGWDQTISNN VQLGGVFTYV RNSNNFDKAS SKN.TLAQVN
HK61IGA     RFSSKSTQTQ LGWDQTISNN VQLGGVFTYV RNSNNFDKAS SKN.TLAQVN
Consensus   ------T--- ----Q----N ---G-VF--- R--N-FD--- ----TL----

                                                                    1850
            1801
Hap         GFAQYQWGDL QF..GVNVGT GISASKMAEE QSRKIHRKAI NYGVNASYQF
HK368IGA    FYSKY.YADN HWYLGIDLGY GKFQSKLQTN HNAKFARHTA QFGLTAGKAF
HK393IGA    FYSKY.YADN HWYLGIDLGY GKFQSKLQTN HNAKFARHTA QFGLTAGKAF
HK715IGA    FYSKY.YADN HWYLGIDLGY GKFQSNLKTN HNAKFARHTA QFGLTAGKAF
HK61IGA     FYSKY.YADN HWYLGIDLGY GKFQSNLQTN NNAKFARHTA QIGLTAGKAF
Consensus   ----Y---D- ----G---G- G---S----- ---K--R--- --G--A---F

                                                                    1900
            1851
Hap         RLGQLGIQPY FGVNRYFIER ENYQSEEVRV KTPSLAFNRY NAGIRVDYTF
HK368IGA    NLGNFGITPI VGVRYSYLSN ADFALDQARI KVNPISVKTA FAQVDLSYTY
HK393IGA    NLGNFGITPI VGVRYSYLSN ADFALDQARI KVNPISVKTA FAQVDLSYTY
HK715IGA    NLGNFGITPI VGVRYSYLSN ANFALAKDRI KVNPISVKTA FAQVDLSYTY
HK61IGA     NLGNFAVKPT VGVRYSYLSN ADFALAQDRI KVNPISVKTA FAQVDLSYTY
Consensus   -LG-----P- -GV------- --------R- K---------- -A-----YT-

                                                                    1950
            1901
Hap         TPTDNISVKP YFFVNYVDVS NANVQTTVNL TVLQQPFGRY WQKEVGLKAE
HK368IGA    .HLGEFSVTP ILSARY.DAN QGSGKINVNG YDFAYNVENQ QQYNAGLKLK
HK393IGA    .HLGEFSVTP ILSARY.DAN QGSGKINVNG YDFAYNVENQ QQYNAGLKLK
HK715IGA    .HLGEFSVTP ILSARY.DTN QGSGKINVNQ YDFAYNVENQ QQYNAGLKLK
HK61IGA     .HLGEFSITP ILSARY.DAN QGNGKINVSV YDFAYNVENQ QQYNAGLKLK
Consensus   ------S--P -----Y-D-- -------V-- ---------- -Q---GLK--

                                                        1982
            1951
Hap         ILHFQISAFI SKSQGSQLGK QQNVGVKLGY RW
HK368IGA    YHNVKLSLIG GLTKAKQAEK QKTAELKLSF SF
HK393IGA    YHNVKLSLIG GLTKAKQAEK QKTAELKLSF SF
HK715IGA    YHNVKLSLIG GLTKAKQAEK QKTAELKLSF SF
HK61IGA     YHNVKLSLIG GLTKAKQAEK QKTAEVKLSF SF
Consensus   ------S--- ------Q--K Q-----KL-- --
```

FIGURE 7H

Fig. 8

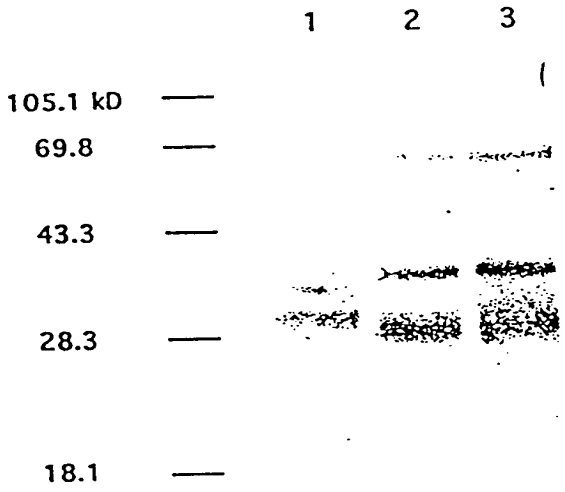


FIGURE 8

Fig. 9

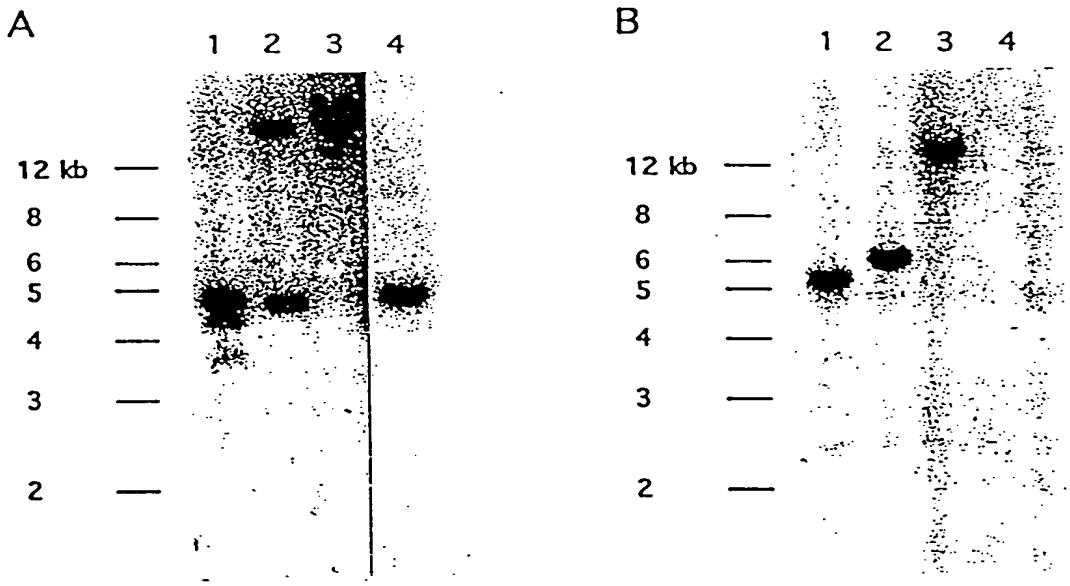


FIGURE 9

Fig. 10

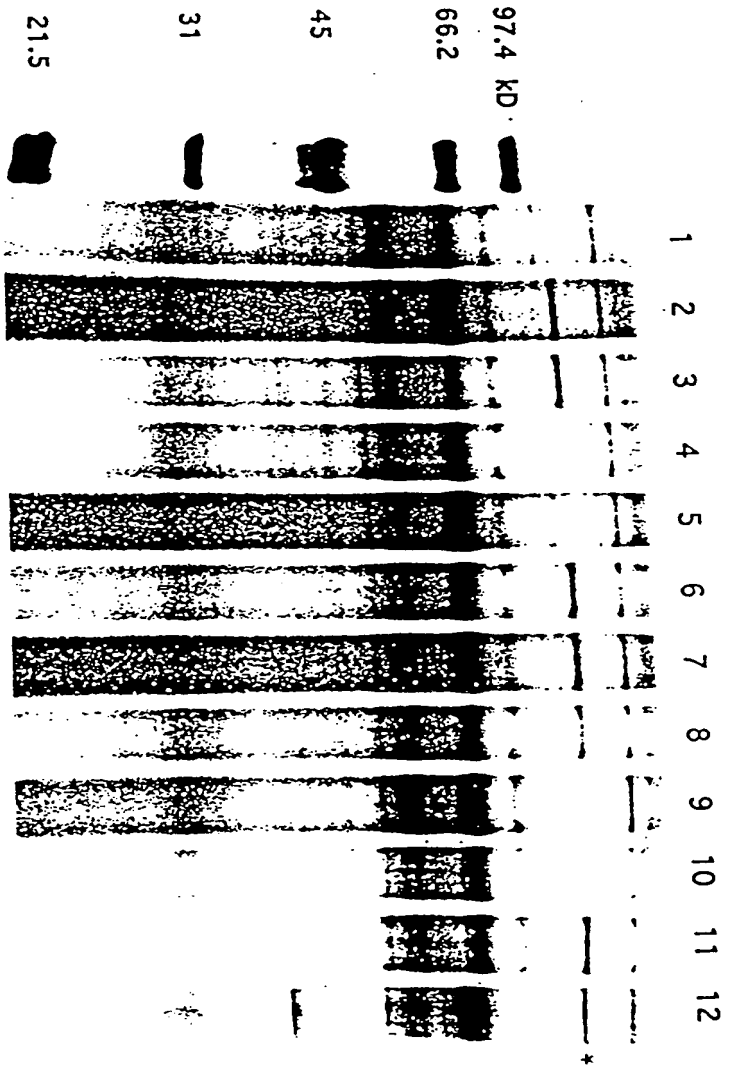


FIGURE 10

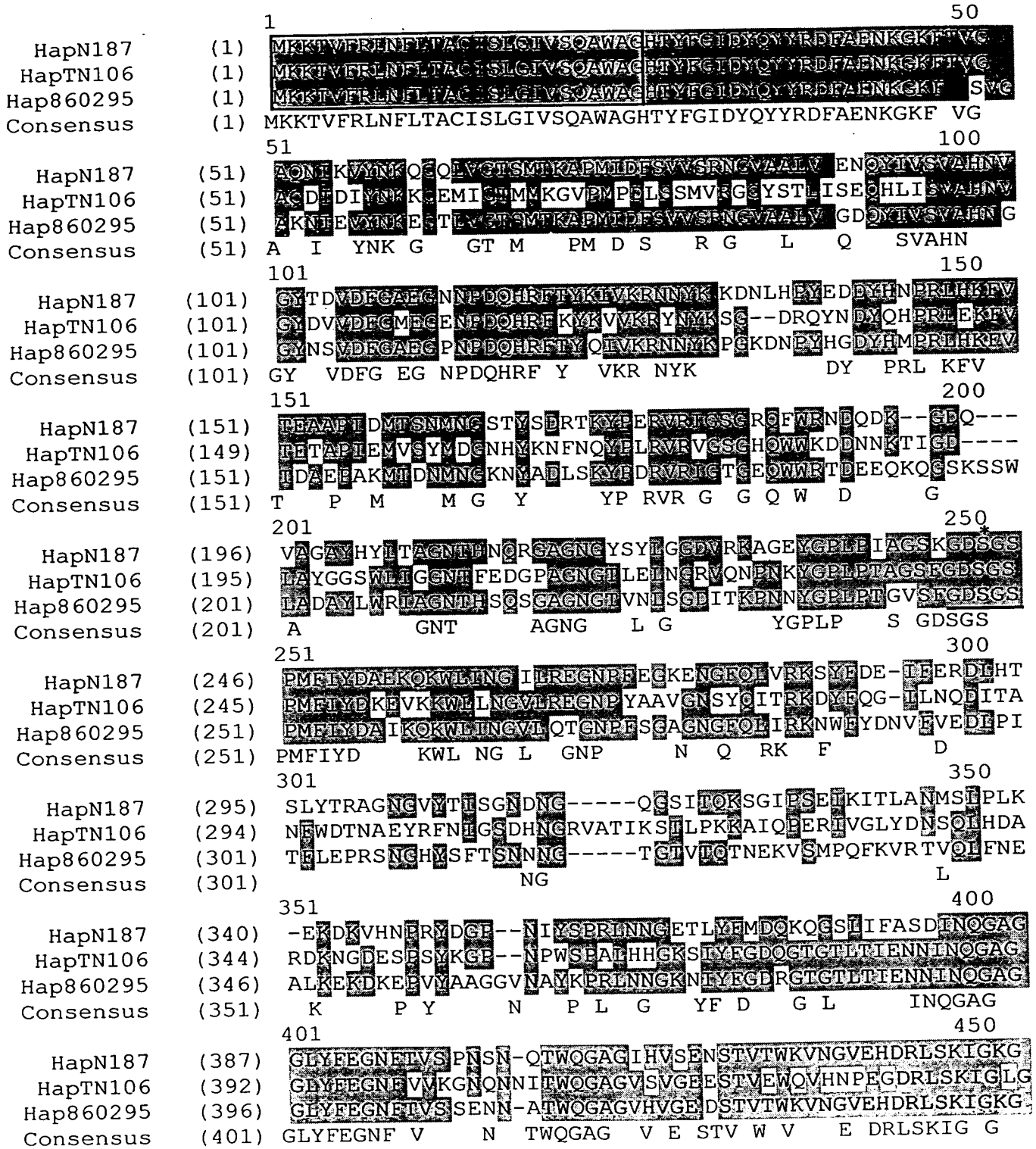

1 50
HapN187 (1) MKKTVFRLNFLTACISLGIVSQAWAGHTYFGIDYQYYRDFAENKGKFTVG
HapTN106 (1) MKKTVFRLNFLTACISLGIVSQAWAGHTYFGIDYQYYRDFAENKGKFTVG
Hap860295 (1) MKKTVFRLNFLTACISLGIVSQAWAGHTYFGIDYQYYRDFAENKGKF SVG
Consensus (1) MKKTVFRLNFLTACISLGIVSQAWAGHTYFGIDYQYYRDFAENKGKF VG
51 100
HapN187 (51) AQNIKVYNKQCQLVGISMTKAPMIDFSVVSRNGVAALV ENQYIVSVAHNV
HapTN106 (51) ACDFDIYNKKCEMICTMMKGVPMPFLSSMVRGGYSTLISEQHLIGVAFNV
Hap860295 (51) AKNIEVYNKESTLVGTSMTKAPMIDFSVVSRNGVAALV GDQYIVSVAHNG
Consensus (51) A I YNK G GT M PM D S R G L Q SVAHN
101 150
HapN187 (101) GYTDVDFGAEGNNPDQHRFTYKLVKRNNYKKDNLHPYEDDYHNPRLHKFV
HapTN106 (101) GYDVVDFGMEGENPDQHRFKYKVVKRYNYKSG--DRQYNDYQHPRLEKFV
Hap860295 (101) GYNSVDFGAEGPNPDQHRFTYQIVKRNNYKPGKDNPYHGDYHMPRLHKFV
Consensus (101) GY VDFG EG NPDQHRF Y VKR NYK DY PRL KFV
151 200
HapN187 (151) TEAAPIDMTSNMNGSTYSDRTKYPERVRIGSGRGFWRNDQDK--GDQ---
HapTN106 (149) TETAPIEMVSYMDGNHYKNFNQYPLRVRVGSGHQWWKDDNNKTIGD----
Hap860295 (151) IDAEPAKMTDNMNGKNYADLSKYPDRVRIGTGEQWWRTDEEQKQGSKSSW
Consensus (151) T P M M G Y YP RVR G G Q W D G
201 250
HapN187 (196) VAGAYHYLTAGNTFNQRGAGNGYSYLGGDVRKAGEYGPLPIAGSKGDSGS
HapTN106 (195) IAYGGSWLIGGNTFEDGPAGNGILELNGRVQNPNKYGPLPTAGSFGDSGS
Hap860295 (201) LADAYLWRIAGNTHSQSGAGNGTVNISGDITKPNNYGPLPTGVSFGDSGS
Consensus (201) A GNT AGNG L G YGPLP S GDSGS
251 300
HapN187 (246) PMFIYDAEKQKWLINGILREGNPFEGKENGFQLVRKSYFDE-IFERDLHT
HapTN106 (245) PMFIYDKEVKKWLLNGVLREGNPYAAVGNSYQITRKDYFQG-LLNQDITA
Hap860295 (251) PMFIYDAIKQKWLINGVLQTGNPFSGAGNGFQLIRKNWFYDNVFVEDIPI
Consensus (251) PMFIYD KWL NG L GNP N Q RK F D
301 350
HapN187 (295) SLYTRAGNGVYTISGNDNG-----QGSITQKSGIPSELIKITLANMSIPLK
HapTN106 (294) NFWDTNAEYRFNIGSDHNGRVATIKSILPKKAIQPERIVGLYDNSQIHDA
Hap860295 (301) TFLEPRSNGHYSFTSNNNG-----TGTVTQTNEKVSMPQFKVRTVQIFNE
Consensus (301) NG L
351 400
HapN187 (340) -EKDKVHNPRYDGP--NIYSPRLNNGETLYFMDQKQGSIIFASDINQGAG
HapTN106 (344) RDKNGDESPSYKGP--NPWSPALHHGKSIYFGDQGTGTLTIENNINQGAG
Hap860295 (346) ALKEKDKEPVYAAGGVNAYKPRLNNGKNIYFGDRGTGTLTIENNINQGAG
Consensus (351) K P Y N P L G YF D G L INQGAG
401 450
HapN187 (387) GLYFEGNFTVSPNSN-QTWQGAGIHVSENSTVTWKVNGVEHDRLSKIGKG
HapTN106 (392) GLYFEGNFVVKGNQNNITWQGAGVSVGEESTVEWQVHNPEGDRLSKIGLG
Hap860295 (396) GLYFEGNFTVSSENN-ATWQGAGVHVGEDSTVTWKVNGVEHDRLSKIGKG
Consensus (401) GLYFEGNF V N TWQGAG V E STV W V E DRLSKIG G

FIG. 11A

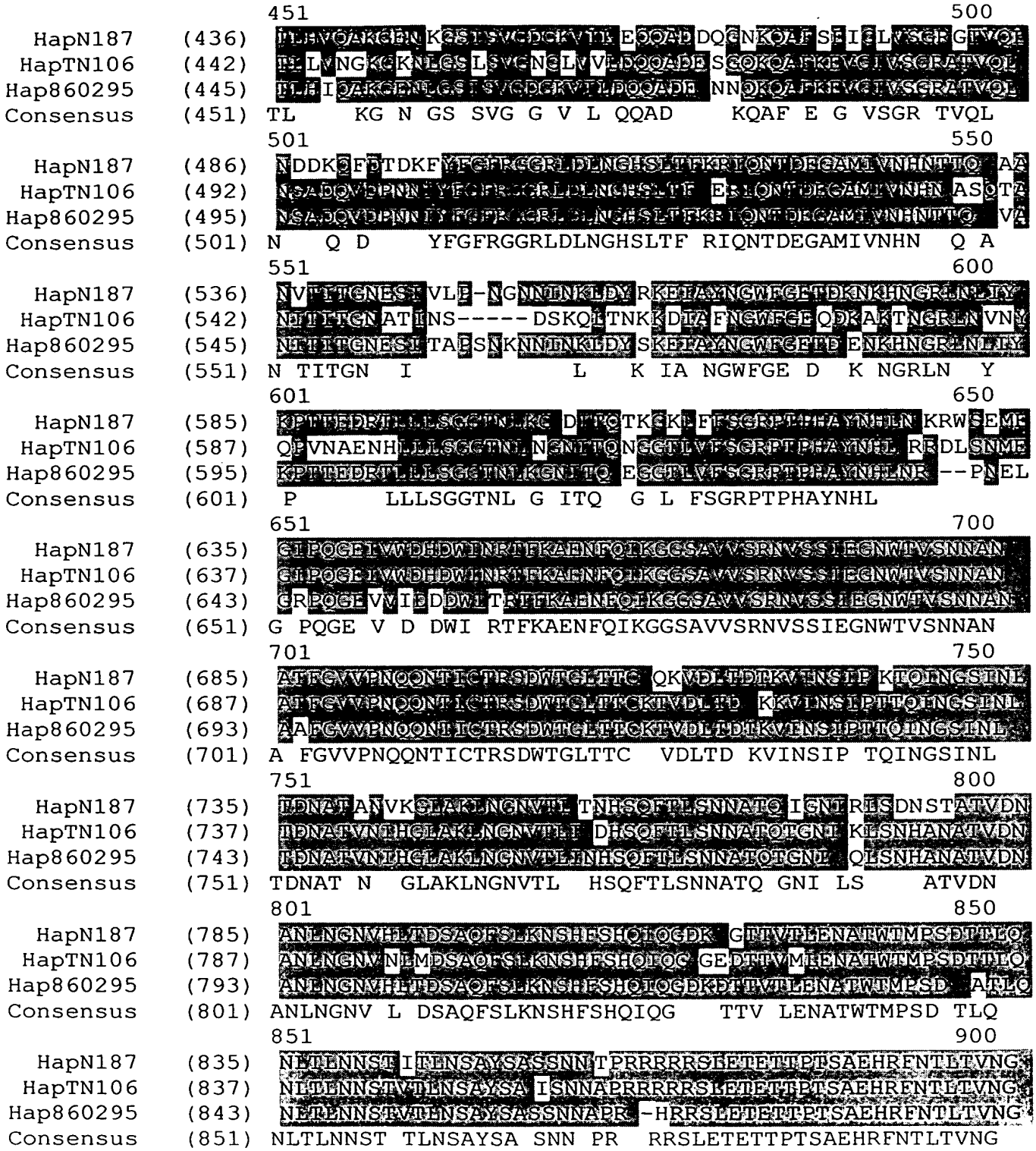

451 500
HapN187 (436)
HapTN106 (442)
Hap860295 (445)
Consensus (451) TL KG N GS SVG G V L QQAD KQAF E G VSGR TVQL
501 550
HapN187 (486)
HapTN106 (492)
Hap860295 (495)
Consensus (501) N Q D YFGFRGGRLDLNGHSLTF RIQNTDEGAMIVNHN Q A
551 600
HapN187 (536)
HapTN106 (542)
Hap860295 (545)
Consensus (551) N TITGN I L K IA NGWFGE D K NGRLN Y
601 650
HapN187 (585)
HapTN106 (587)
Hap860295 (595)
Consensus (601) P LLLSGGTNL G ITQ G L FSGRPTPHAYNHL
651 700
HapN187 (635)
HapTN106 (637)
Hap860295 (643)
Consensus (651) G PQGE V D DWI RTFKAENFQIKGGSAVVSRNVSSIEGNWTVSNNAN
701 750
HapN187 (685)
HapTN106 (687)
Hap860295 (693)
Consensus (701) A FGVVPNQQNTICTRSDWTGLTTC VDLTD KVINSIP TQINGSINL
751 800
HapN187 (735)
HapTN106 (737)
Hap860295 (743)
Consensus (751) TDNAT N GLAKLNGNVTL HSQFTLSNNATQ GNI LS ATVDN
801 850
HapN187 (785)
HapTN106 (787)
Hap860295 (793)
Consensus (801) ANLNGNV L DSAQFSLKNSHFSHQIQG TTV LENATWTMPSD TLQ
851 900
HapN187 (835)
HapTN106 (837)
Hap860295 (843)
Consensus (851) NLTLNNST TLNSAYSA SNN PR RRSLETETTPTSAEHRFNTLTVNG

FIG. 11B

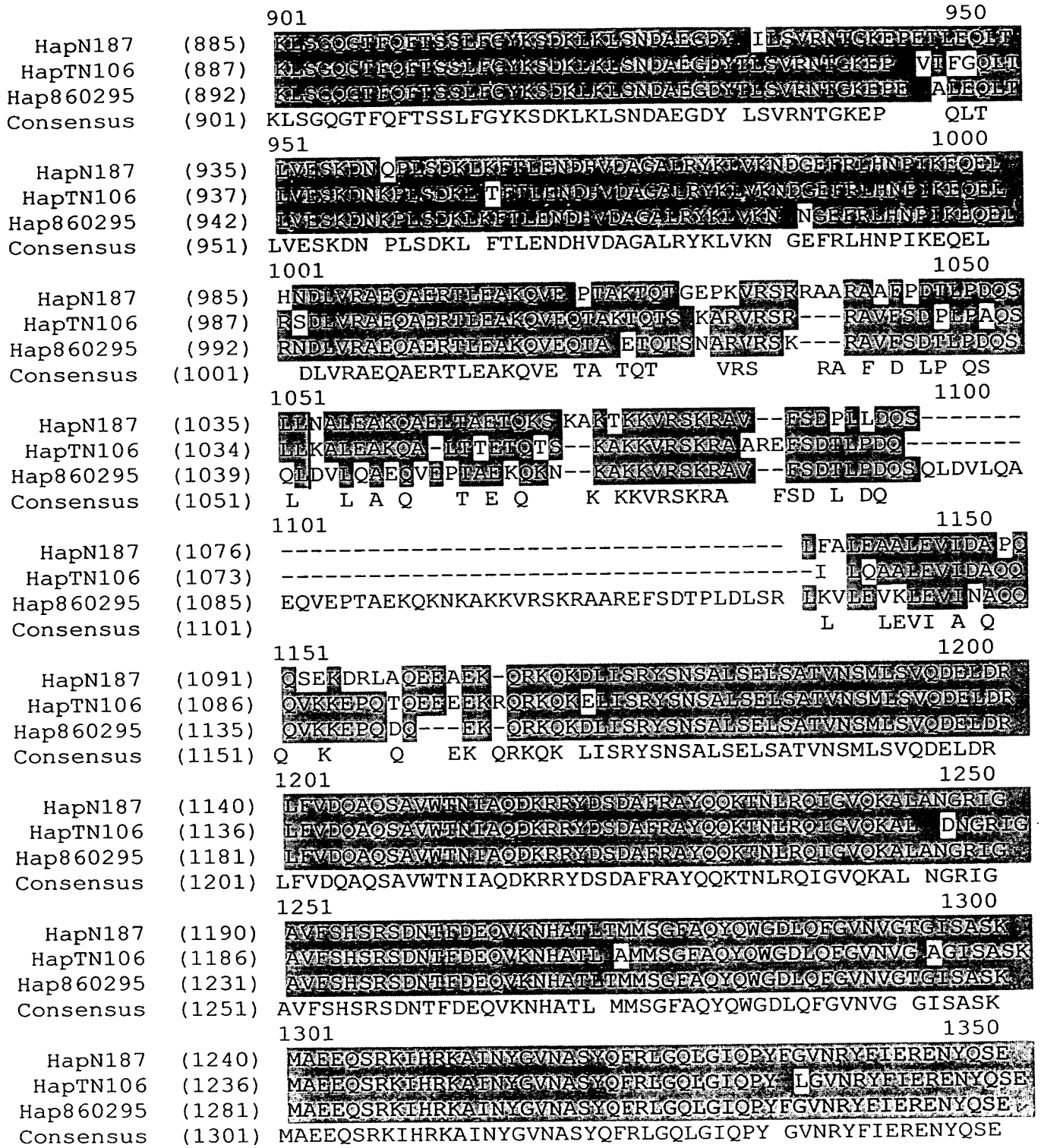

901 950
HapN187 (885)
HapTN106 (887)
Hap860295 (892)
Consensus (901) KLSGQGTFQFTSSLFGYKSDKLKLSNDAEGDY LSVRNTGKEP QLT
951 1000
HapN187 (935)
HapTN106 (937)
Hap860295 (942)
Consensus (951) LVESKDN PLSDKL FTLENDHVDAGALRYKLVKN GEFRLHNPIKEQEL
1001 1050
HapN187 (985)
HapTN106 (987)
Hap860295 (992)
Consensus (1001) DLVRAEQAERTLEAKQVE TA TQT VRS RA F D LP QS
1051 1100
HapN187 (1035)
HapTN106 (1034)
Hap860295 (1039)
Consensus (1051) L L A Q T E Q K KKVRSKRA FSD L DQ
1101 1150
HapN187 (1076)
HapTN106 (1073)
Hap860295 (1085) EQVEPTAEKQKNKAKKVRSKRAAREFSDTPLDLSR LKVLEVKLEVINAQQ
Consensus (1101) L LEVI A Q
1151 1200
HapN187 (1091)
HapTN106 (1086)
Hap860295 (1135)
Consensus (1151) Q K Q EK QRKQK LISRYSNSALSELSATVNSMLSVQDELDR
1201 1250
HapN187 (1140)
HapTN106 (1136)
Hap860295 (1181)
Consensus (1201) LFVDQAQSAVWTNIAQDKRRYDSDAFRAYQQKTNLRQIGVQKAL NGRIG
1251 1300
HapN187 (1190)
HapTN106 (1186)
Hap860295 (1231)
Consensus (1251) AVFSHSRSDNTFDEQVKNHATL MMSGFAQYQWGDLQFGVNVG GISASK
1301 1350
HapN187 (1240)
HapTN106 (1236)
Hap860295 (1281)
Consensus (1301) MAEEQSRKIHRKAINYGVNASYQFRLGQLGIQPY GVNRYFIERENYQSE

FIG. 11C

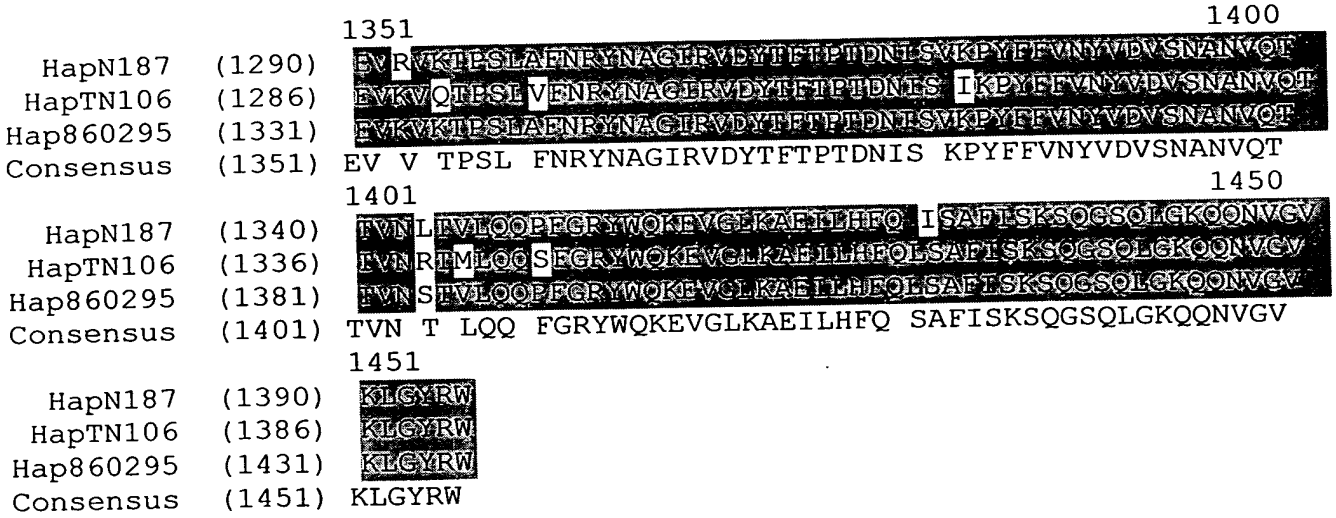

1351
1400
HapN187 (1290) EVRVKTPSLAFNRYNAGIRVDYTFTPTDNLSVKPYFFVNYVDVSNANVQT
HapTN106 (1286) EVKVQTPSLVENRYNAGIRVDYTFTPTDNIS IKPYFFVNYVDVSNANVQT
Hap860295 (1331) EVKVKTPSLAENRYNAGIRVDYTFTPTDNLSVKPYFFVNYVDVSNANVQT
Consensus (1351) EV V TPSL FNRYNAGIRVDYTFTPTDNIS KPYFFVNYVDVSNANVQT
1401
1450
HapN187 (1340) TVNLEVLQQPEGRYWQKEVGLKAEILHFQ ISAFISKSQGSQLGKQQNVGV
HapTN106 (1336) TVNRTMLQQSEGRYWQKEVGLKAEILHFQLSAFISKSQGSQLGKQQNVGV
Hap860295 (1381) TVNSTVLQQPEGRYWQKEVGLKAEILHFQLSAFISKSQGSQLGKQQNVGV
Consensus (1401) TVN T LQQ FGRYWQKEVGLKAEILHFQ SAFISKSQGSQLGKQQNVGV
1451
HapN187 (1390) KLGYRW
HapTN106 (1386) KLGYRW
Hap860295 (1431) KLGYRW
Consensus (1451) KLGYRW

FIG. 11D

Fig. 12

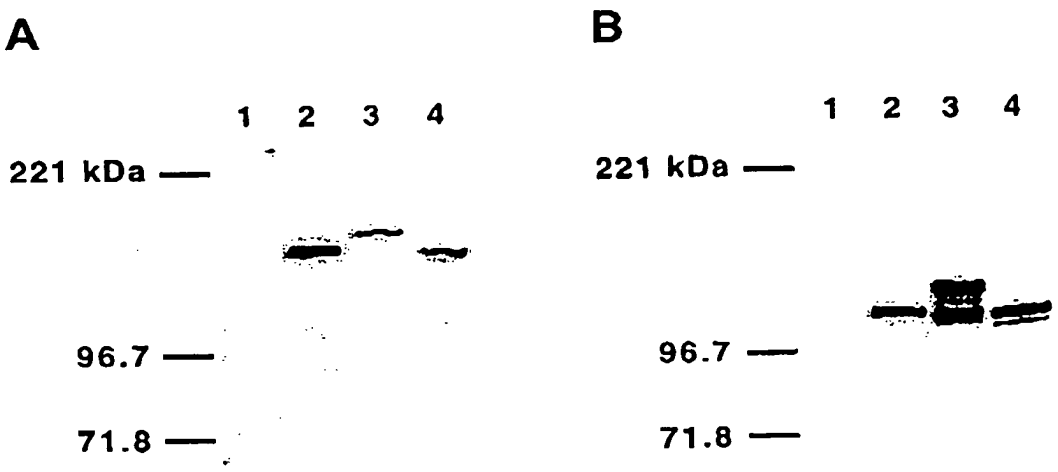


Fig B

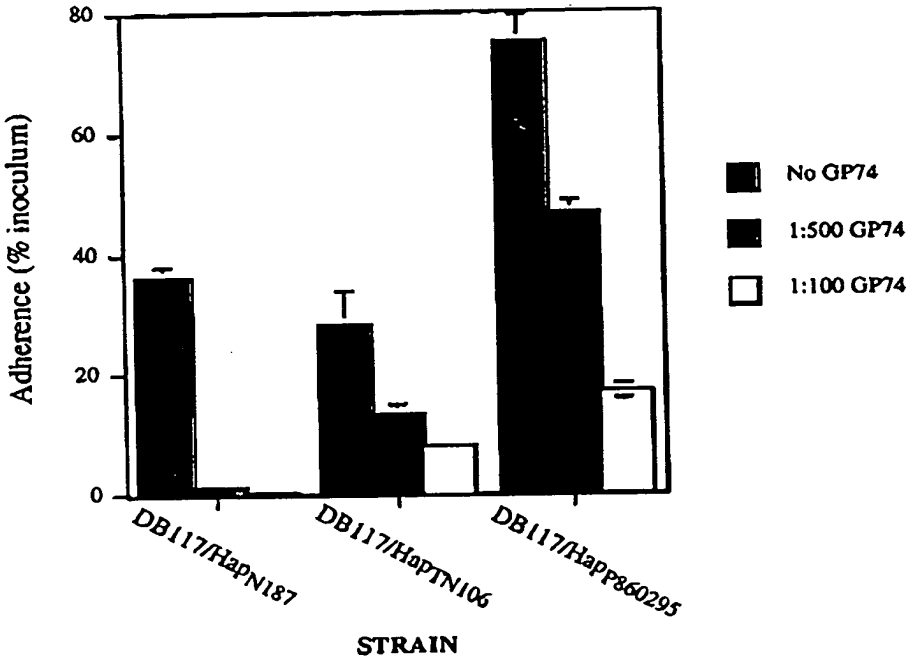

80
60
40
20
0
Adherence (% inoculum)
No GP74
1:500 GP74
1:100 GP74
DB117/HapN187
DB117/HapTN106
DB117/HapP860295
STRAIN

Fig 14

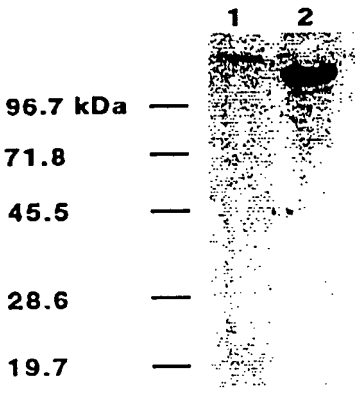
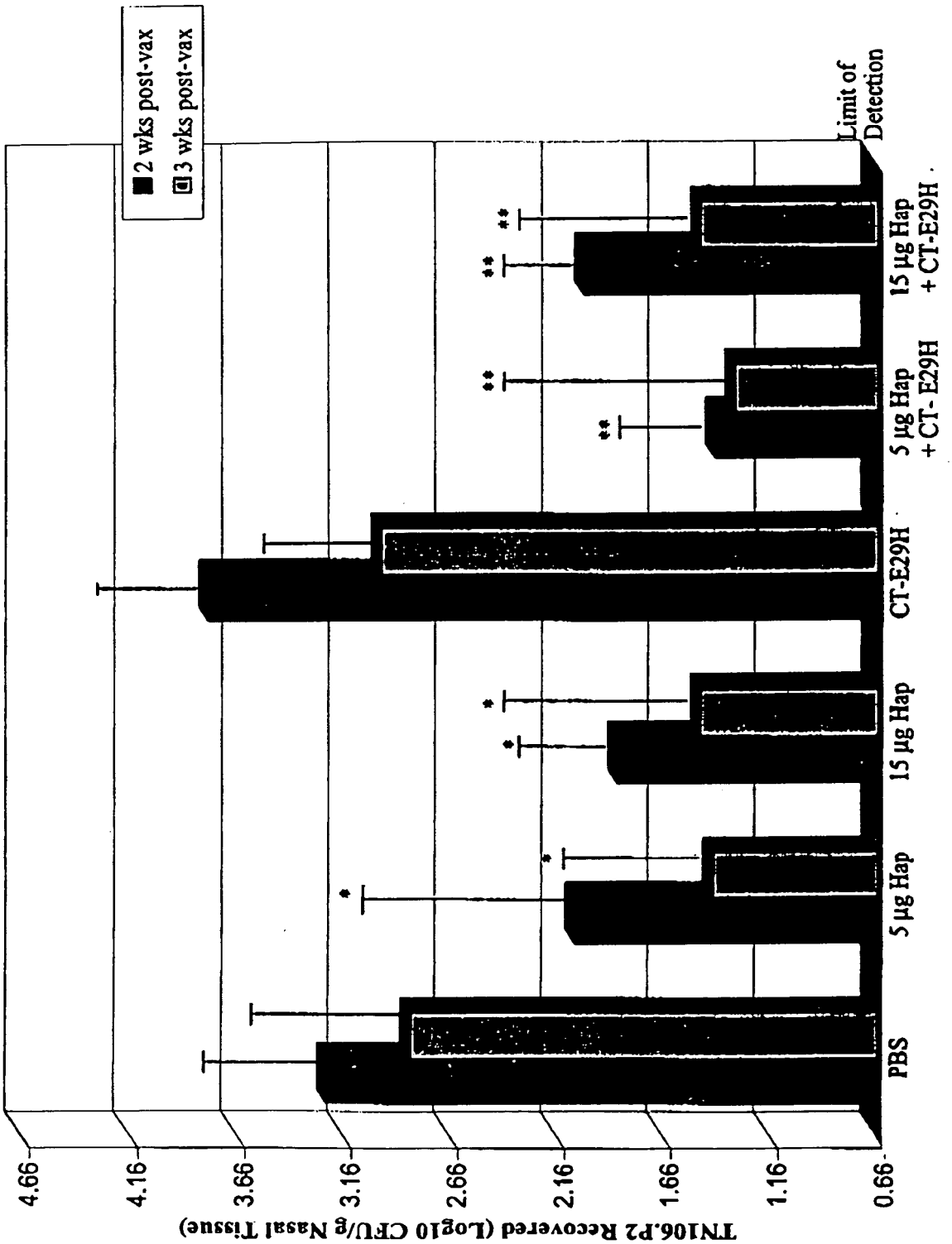


Fig 15


2 wks post-vax
3 wks post-vax
TN106.P2 Recovered (Log10 CFU/g Nasal Tissue)
4.66
4.16
3.66
3.16
2.66
2.16
1.66
1.16
0.66
PBS
5 µg Hap
15 µg Hap
CT-E29H
5 µg Hap + CT- E29H
15 µg Hap + CT-E29H
Limit of Detection

Nucleotide sequence for NTHi strain 11 *hap* gene (start codon to stop codon):

```
   1  ATGAAAAAAA CTGTATTTCG TCTTAATTTT TTAACCGCTT GCATTTCATT
  51  AGGGATAGTA TCGCAAGCGT GGGCAGGTCA TACTTATTTT GGGATTGACT
 101  ACCAATATTA TCGTGATTTT GCCGAGAATG AAGGCAAGTT TGCAGTTGGG
 151  GCTAAAAATA TTGATGTTTA TAACAAAGAA GGGCAATTAG TTGGCACATC
 201  AATGACAAAA GCCCCGATGA TTGATTTCTC AGTCGTTTCC AGAAATGGAG
 251  TTGCTGCCTT AGTAGGCGAT CAGTATATTG TGAGTGTGGC ACATAATGTA
 301  GGCTATACCA ATGTGGATTT TGGTGCTGAA GGACAAAATC CTGATCAACA
 351  TCGTTTTACT TATAAAATTG TGAAACGGAA TAATTATAAT CACGATGCGA
 401  AGCACCGCTA TCTAGATGAC TACCATAATC CACGTTTACA TAAATTTGTA
 451  ACGGATGCGG CACCAATTGA TATGACTTCA CATATGGATG GCAATAAGTA
 501  TGCAAATAAG GAAAAATATC CTGAACGAGT ACGCGTCGGA TCTGGAGATC
 551  AGTATTGGGA TGACGATCAA AACAACAGAA CTTATTTATC TGACGGATAT
 601  AATTATTTAA CAGGTGGGAA TACATATAAT CAAAGCGGTA GAGGTGATGG
 651  ATATTCATAT GTGAGAGGTG ATATTCGCAA AGTTGGCGAT TATGGTCCAT
 701  TACCGATTGC AAGTTCATTC GGGGACAGTG GATCTCCAAT GTTTATTTAT
 751  GATGCTGAAA CACAAAAATG gcTAATTAAT GGAGTATTGC GGGAGGGGCA
 801  ACCTTATACA GGCGAATTCG ATGGATTTCA ATTAGCCCGT AAATCTTTCC
 851  TTGATGAAAT TATACGCAAA GATCAACCAA ATGGTTTTTT AACCCCTAAG
 901  GGGAATGGCG TTTATACCAT TTCTAAAAGT GACGATGGGA TAGGAGTTGT
 951  TACTTCGAAA ATTGGAAAAC CTCGTGAAAT ACCTTTAGCG AACAACAAAT
1001  TAAAAATAGA AGATAAAGAT ACTGTCTATA ATAACAGATA TAATGGTCCT
1051  AATATTTATT CTCCTCAATT AAACAATGGC AAGAATATTT ATTTTGGAGA
1101  TGAAGAATTA GGATCCATAA CTTTAACGAC TGATATCGAT CAAGGTGCAG
1151  GCGGTTTGTA TTTTGAGGGG GATTTTATAG TTTCGCCTAC CAAAAATGAA
1201  ACGTGGAAAG GTGCGGGCAT TCATGTCAGT GAAATTAGTA CCGTTACTTG
1251  GAAAGTAAAC GGCGTGGAAA ATGATCGACT TTCTAAAATC GGTAAAGGAA
1301  CATTACACGT TAAAGCCAAA GGGGAAAATA AAGGTTCGAT CAGCGTAGGC
1351  GATGGTAAAG TCATTTTGGA GCAGCAGGCA GACGATCAAG GCAACAAACA
1401  AGCCTTTAGT GAAATTGGCT TGGTTAGCGG CAGAGGGACT GTTCAATTAA
1451  ACGATGATAA ACAATTTGAT ACCGATAAAT TTTATTTCGG CTTTCGTGGT
1501  GGTCGCTTAG ATCTTAACGG ACATTCATTA ACCTTTAAAC GTATCCAAAA
1551  TACGGACGAG GGGGCGATGA TTGTGAACCA TAATACAACT CAAGTCGCTA
1601  ATATTACTAT TACTGGGAAC GAAAGTATTA CTGCTCCATC TAATAAAAAT
1651  AATATTAATA AACTTGATTA CAGCAAAGAA ATTGCCTACA ACGGCTGGTT
1701  TNGCGAAACA GATAAAAATA AACATAATGG ACGATTAAAC CTTATTTATA
1751  AACCAACCAC AGAAGATCGT ACTTTGCTAC TTTCAGGCGG CACAAACTTA
1801  AAAGGCGATA TTACTCAAAC AAAAGGTAAA CTATTTTTCA GCGGTAGACC
1851  GACACCCCAC GCCTACAATC ATTTAGACAA ACGTTGGTCA GAAATGGAAG
1901  GTATCCCACA AGGCGAAATT GTGTGGGATT ACGATTGGAT TAACCGCACA
1951  TTTAAAGCTG AAAACTTCCA AATTAAAGGC GGAAGTGCGG TGGTTTCTCG
2001  CAATGTTTCT TCAATTGAGG GAAATTGGAC AGTCAGCAAT AATGCAAATG
```

Fig. 16A

2051 CCACATTTGG TGTTGTGCCA AATCAGCAAA ATACCATTTG CACGCGTTCA
2101 GATTGGACAG GATTAACGAC TTGTAAAACA GTTAATTTAA CCGATAAAAA
2151 AGTTATTGAT TCCATACCGA CAACACAAAT TAATGGTTCT ATTAATTTAA
2201 CTGATAATGC AACAGTGAAT ATTAATGGTT TAGCAAAACT TAATGGTAAT
2251 GTCACTTTAA TAAATCATAG CCAATTTACA TTGAGCAACA ATGCCACCCA
2301 AATAGGCAAT ATCAAACTTT CAAATCACGC AAATGCAAGG GTAAATAATG
2351 CCACTTTAAT GGGCGATGTG AATTTAGCGG ATACTAGCCG TTTTACATTA
2401 AGCAATCAAG CAACACAGAT TGGCACAATC AGTCTTCATC AGCAAGCTCA
2451 AGCAACAGTG GATAATGCAA ACTTGAACGG TAATGTGCAT TTAACGGATT
2501 CTGCCAGATT TTCTTTAAAA AACAGTCATT TTTCGCACCA AATTCAGGGC
2551 GACAAAGACA CAACAGTGAC GTTGGAAAAT GCGACTTGGA CAATGCCTAG
2601 CGATACTACA TTGCAGAATT TAACGCTAAA TAATAGTACT GTTACGTTAA
2651 ATTCAGCTTA TTCAGCTAGC TCAAATAATG CGCCACGTCG CCgCCGTTCA
2701 TTAGAGACGG AAACAACGCC AACATCGGCA GAACATCGTT TCAACACATT
2751 GACAGTAAAT GGTAAATTGA GCGGGCAAGG CACATTCCAA TTTACTCCAT
2801 CTTTATTTGG CTATGAAAGC GATAAATTAA AATTATCCAA TGACGCTGAG
2851 GGCGATTACA CATTATCTGT TCGCAACACA GGCAAAGAAC CCGTGACCCT
2901 TGAGCAATTA ACTTTGGTTG AAAGCAAAGA TAATAAACCG TTATCAGACA
2951 AACTCAAATT TACTTTAGAA AATGACCACG TTGATGCAGG TGCATTACGT
3001 TATAAATTAG TGAAGAATAA GGGCGAATTC CGCTTGCATA ACCCAATAAA
3051 AGAGCAGGAA TTGCGCTCTG ATTTAGTAAG AGCAGAGCAA GCAGAACGAA
3101 CATTAGAAGC CAAACAAGTT GAACAGACTG CTGAAACACA AACAAGTAAT
3151 GCAAGAGTGC GGTCAAGAAG AGCGGTGTTG TCTGATACCC CGTCTGCTCA
3201 AAGCCTGTTA AACGCATTAG AAGTCAAACA AGCTGAACCG AATGCTAAAA
3251 CACAAAAAAG TAAGGCAAAA ACAAAAAAAG CGCGGTCAAA AAGAGCATTG
3301 AGAGAAGCGT TTTCTGATAC CCCGCCTGAT CTAAGCCAGT TAAACGTATT
3351 AGAAGCCGCA CTTAAGGTTA TTAATGCCCA ACCGCAAACA GAAAAAGAAC
3401 GTCAAGCTCA AGAGGAAGAA GCGAAAAGAC AACGCaAACA AAAAGACTTG
3451 ATCAGCCGTT ACTCAAATAG TGCGTTATCG GAGTTGTCTG CAACAGTAAA
3501 TAGTATGCTT TCCGTTCAAG ATGAATTGGA TCGTCTTTTT GTAGATCAAG
3551 CACAATCTGC CCTGTGGACA AATATCGCAC AGGATAAAAG ACGCTATGAT
3601 TCTGATGCGT TCCGTGCTTA TCAGCAGAAA ACGAACTTGC GTCAAATTGG
3651 GGTGCAAAAA GCCTTAGATA ATGGACGAAT TGGGGCGGTT TTCTCGCATA
3701 GCCGTTCAGA TAATACCTTT GACGAACAGG TTAAAAATCA CGCGACATTA
3751 ACGATGATGT CGGGTTTTGC CCAATATCAA TGGGGCGATT TACAATTTGG
3801 TGTAAACGTG GGCGCGGGAA TTAGTGCGAG TAAAATGGCT GAAGAACAAA
3851 GCCGAAAAAT TCATCGAAAA GCGATAAATT ATGGTGTGAA TGCAAGTTAT
3901 CAGTTCCGTT TAGGGCAATT GGGTATTCAG CCTTATTTGG GTGTTAATCG
3951 ATATTTTATT GAACGTGAAA ATTATCAATC TGAAGAAGTG AAAGTGCAAA
4001 CACCGAGCCT TGCATTTAAT CGCTATAATG CTGGCATTCG AGTTGATTAT
4051 ACATTTACCC CGACAGATAA TATCAGCGTT AAGCCTTATT TCTTTGTCAA
4101 TTATGTTGAT GTTTCAAACG CTAACGTACA AACCACTGTA AATAGCACGA
4151 TGTTGCAACA ATCATTTGGG CGTTATTGGC AAAAAGAAGT GGGATTAAAG
4201 GCAGAAATTT TACATTTCCA ACTTTCCGCT TTTATCTCAA AATCTCAAGG

Fig. 16B

```
4251  TTCACAACTC GGTAAACAGC AAAATGTGGG CGTGAAATTG GGCTATCGTT
4301  GGTAA
```

Fig. 16C

Amino acid sequence for NTHi strain 11 Hap protein (first amino acid to last amino acid):

```
   1  MKKTVFRLNF LTACISLGIV SQAWAGHTYF GIDYQYYRDF AENEGKFAVG
  51  AKNIDVYNKE GQLVGTSMTK APMIDFSVVS RNGVAALVGD QYIVSVAHNV
 101  GYTNVDFGAE GQNPDQHRFT YKIVKRNNYN HDAKHRYLDD YHNPRLHKFV
 151  TDAAPIDMTS HMDGNKYANK EKYPERVRVG SGDQYWDDDQ NNRTYLSDGY
 201  NYLTGGNTYN QSGRGDGYSY VRGDIRKVGD YGPLPIASSF GDSGSPMFIY
 251  DAETQKWLIN GVLREGQPYT GEFDGFQLAR KSFLDEIIRK DQPNGFLTPK
 301  GNGVYTISKS DDGIGVVTSK IGKPREIPLA NNKLKIEDKD TVYNNRYNGP
 351  NIYSPQLNNG KNIYFGDEEL GSITLTTDID QGAGGLYFEG DFIVSPTKNE
 401  TWKGAGIHVS EISTVTWKVN GVENDRLSKI GKGTLHVKAK GENKGSISVG
 451  DGKVILEQQA DDQGNKQAFS EIGLVSGRGT VQLNDDKQFD TDKFYFGFRG
 501  GRLDLNGHSL TFKRIQNTDE GAMIVNHNTT QVANITITGN ESITAPSNKN
 551  NINKLDYSKE IAYNGWFXET DKNKHNGRLN LIYKPTTEDR TLLLSGGTNL
 601  KGDITQTKGK LFFSGRPTPH AYNHLDKRWS EMEGIPQGEI VWDYDWINRT
 651  FKAENFQIKG GSAVVSRNVS SIEGNWTVSN NANATFGVVP NQQNTICTRS
 701  DWTGLTTCKT VNLTDKKVID SIPTTQINGS INLTDNATVN INGLAKLNGN
 751  VTLINHSQFT LSNNATQIGN IKLSNHANAR VNNATLMGDV NLADTSRFTL
 801  SNQATQIGTI SLHQQAQATV DNANLNGNVH LTDSARFSLK NSHFSHQIQG
 851  DKDTTVTLEN ATWTMPSDTT LQNLTLNNST VTLNSAYSAS SNNAPRRRRS
 901  LETETTPTSA EHRFNTLTVN GKLSGQGTFQ FTPSLFGYES DKLKLSNDAE
 951  GDYTLSVRNT GKEPVTLEQL TLVESKDNKP LSDKLKFTLE NDHVDAGALR
1001   YKLVKNKGEF RLHNPIKEQE LRSDLVRAEQ AERTLEAKQV EQTAETQTSN
1051   ARVRSRRAVL SDTPSAQSLL NALEVKQAEP NAKTQKSKAK TKKARSKRAL
1101   REAFSDTPPD LSQLNVLEAA LKVINAQPQT EKERQAQEEE AKRQRKQKDL
1151   ISRYSNSALS ELSATVNSML SVQDELDRLF VDQAQSALWT NIAQDKRRYD
1201   SDAFRAYQQK TNLRQIGVQK ALDNGRIGAV FSHSRSDNTF DEQVKNHATL
1251   TMMSGFAQYQ WGDLQFGVNV GAGISASKMA EEQSRKIHRK AINYGVNASY
1301   QFRLGQLGIQ PYLGVNRYFI ERENYQSEEV KVQTPSLAFN RYNAGIRVDY
1351   TFTPTDNISV KPYFFVNYVD VSNANVQTTV NSTMLQQSFG RYWQKEVGLK
1401    AEILHFQLSA FISKSQGSQL GKQQNVGVKL GYRW
```

Fig. 17

Nucleotide sequence for NTHi strain TN106 *hap* gene (start codon begins at position 422, stop codon begins at position 4595):

```
   1  TGGCGGCGGA CAAATTATTG CGACGGGTAC ACCAGAACAA GTTGCTAAAG
  51  TAAAAAGTTC CCACACCGCT CGCTTCCTTA AACCGATTTT AGAAAAACCT
 101  TAGAAAAAAT GACCGCACTT TCAGAGAAAA CTCACATAAA GTGCGGTTAT
 151  TTTATTAGTG ATATTGTTTT AATTTTAGTT ATCTGTATAA ATTACATACA
 201  ATATTAATCC ATCGCAAGAT TAGATTACCC ACTAAGTATT AAGCAAAAAC
 251  CTAGAAATTT TGGCTTAATT ACTATATAGT TTTACTCATT TATTTTCTTT
 301  TGTGCCTTTT AGTTCATTTT TTTAGCTGAA ATCCCTTAGA AAATCACCGC
 351  ACTTTTATTG TTCAATAGTC GTTTAACCAC GTATTTTTTA ATACGAAAAA
 401  TTACTTAATT AAATAAACAT TATGAAAAAA ACTGTATTTC GTCTGAATTT
 451  TTTAACCGCT TGCATTTCAT TAGGGATAGT ATCGCAAGCG TGGGCAGGTC
 501  ATACTTATTT TGGGATTGAC TACCAATATT ATCGTGATTT TGCCGAGAAT
 551  AAAGGGAAGT TTACAGTTGG GGCTCAAGAT ATTGATATCT ACAATAAAAA
 601  AGGGGAAATG ATAGGTACGA TGATGAAAGG TGTGCCTATG CCTGATTTAT
 651  CTTCCATGGT TCGTGGTGGT TATTCAACAT TGATAAGTGA GCAGCATTTA
 701  ATTAGCGTCG CACATAATGT AGGGTATGAT GTCGTTGATT TTGGTATGGA
 751  GGGGGAAAAT CCAGACCAAC ATCGTTTTAA GTATAAAGTT GTTAAACGAT
 801  ATAATTATAA GAGCGGTGAT AGACAATATA ATGATTATCA ACATCCAAGA
 851  TTAGAGAAAT TTGTAACGGA AACTGCACCT ATTGAAATGG TTTCATATAT
 901  GGATGGTAAT CATTACAAAA ATTTTAATCA ATATCCTTTG CGAGTTAGAG
 951  TTGGAAGTGG GCATCAATGG TGGAAAGACG ATAATAATAA AACCATTGGA
1001  GACTTAGCCT ATGGAGGTTC ATGGTTAATA GGTGGAAATA CCTTTGAAGA
1051  TGGACCAGCT GGTAACGGTA CATTAGAATT AAATGGGCGA GTACAAAATC
1101  CTAATAAATA TGGTCCACTA CCTACGGCAG GTTCATTCGG GGATAGTGGT
1151  TCTCCAATGT TTATTTATGA TAAGGAAGTT AAGAAATGGT TATTAAATGG
1201  CGTGTTACGT GAAGGAAATC CTTATGCTGC AGTAGGAAAC AGCTATCAAA
1251  TTACACGAAA AGATTATTTT CAAGGTATTC TTAATCAAGA CATTACAGCT
1301  AATTTTTGGG ATACTAATGC TGAATATAGA TTTAATATAG GGAGTGACCA
1351  CAATGGAAGA GTGGCAACAA TCAAAAGTAC ATTACCTAAA AAAGCTATTC
1401  AGCCTGAACG AATAGTGGGT CTTTATGATA ATAGCCAACT TCATGATGCT
1451  AGAGATAAAA ATGGCGATGA ATCTCCCTCT TATAAAGGTC CTAATCCATG
1501  GTCGCCAGCA TTACATCATG GGAAAAGTAT TTACTTTGGC GATCAAGGAA
1551  CAGGAACTTT AACAATTGAA AATAATATAA ATCAAGGTGC AGGTGGATTG
1601  TATTTTGAAG GTAATTTTGT TGTAAAAGGC AATCAAAATA ATATAACTTG
1651  GCAAGGTGCA GGCGTTTCTG TTGGAGAAGA AAGTACTGTT GAATGGCAGG
1701  TGCATAATCC AGAAGGCGAT CGCTTATCCA AAATTGGGCT GGGAACCTTA
1751  CTTGTTAATG GTAAAGGGAA AAACTTAGGA AGCCTGAGTG TCGGTAACGG
1801  TTTGGTTGTG TTAGATCAAC AAGCAGATGA ATCAGGTCAA AAACAAGCCT
1851  TTAAAGAAGT TGGCATTGTA AGTGGTAGAG CTACCGTTCA ACTAAATAGT
1901  GCAGATCAAG TTGATCCTAA CAATATTTAT TTCGGCTTTC GTGGTGGTCG
1951  CTTAGATCTT AATGGGCATT CATTAACCTT TGAACGTATC CAAAATACGG
2001  ATGAAGGCGC GATGATTGTG AACCACAACG CTTCTCAAAC CGCAAATATT
```

Fig. 18A

```
2051  ACGATTACAG GCAACGCAAC TATTAATTCA GATAGCAAAC AACTTACTAA
2101  TAAAAAAGAT ATTGCATTTA ACGGCTGGTT TGGTGAGCAA GATAAAGCTA
2151  AAACAAATGG TCGTTTAAAT GTGAATTATC AACCAGTTAA TGCAGAAAAT
2201  CATTTGTTGC TTTCTGGGGG GACAAATTTA AACGGCAATA TCACGCAAAA
2251  TGGTGGTACG TTAGTTTTTA GTGGTCGTCC AACGCCTCAT GCTTACAATC
2301  ATTTAAGAAG AGACTTGTCT AACATGGAAG GTATCCCACA AGGCGAAATT
2351  GTGTGGGATC ACGATTGGAT CAACCGCACA TTTAAAGCTG AAAACTTCCA
2401  AATTAAAGGC GGAAGTGCGG TGGTTTCTCG CAATGTTTCT TCAATTGAGG
2451  GAAATTGGAC AGTCAGCAAT AATGCAAATG CCACATTTGG TGTTGTGCCA
2501  AATCAGCAAA ATACCATTTG CACGCGTTCA GATTGGACAG GATTAACGAC
2551  TTGTAAAACA GTTGATTTAA CCGATAAAAA AGTTATTAAT TCCATACCGA
2601  CAACACAAAT TAATGGTTCT ATTAATTTAA CTGATAATGC AACAGTGAAT
2651  ATTCATGGTT TAGCAAAACT TAATGGTAAT GTCACTTTAA TAGATCACAG
2701  CCAATTTACA TTGAGCAACA ATGCCACCCA AACAGGCAAT ATCAAACTTT
2751  CAAATCACGC AAATGCAACG GTGGACAATG CAAATTTGAA CGGTAATGTG
2801  AATTTAATGG ATTCTGCTCA ATTTTCTTTA AAAAACAGCC ATTTTTCGCA
2851  CCAAATCCAA GGTGGGGAAG ACACAACAGT GATGTTGGAA AATGCGACTT
2901  GGACAATGCC TAGCGATACC ACATTGCAGA ATTTAACGCT AAATAATAGT
2951  ACTGTTACGT TAAATTCAGC TTATTCAGCT ATCTCAAATA ATGCGCCACG
3001  CCGTCGCCGC CGTTCATTAG AGACGGAAAC AACGCCAACA TCGGCAGAAC
3051  ATCGTTTCAA CACATTGACA GTAAATGGTA AATTGAGCGG GCAAGGCACA
3101  TTCCAATTTA CTTCATCTTT ATTTGGCTAT AAAAGCGATA AATTAAAATT
3151  ATCCAATGAC GCTGAGGGCG ATTACACATT ATCTGTTCGC AACACAGGCA
3201  AAGAACCCGT GACCTTTGGG CAATTAACTT TGGTTGAAAG CAAAGATAAT
3251  AAACCGTTAT CAGACAAACT CACATTCACG TTAGAAAATG ACCACGTTGA
3301  TGCAGGTGCA TTACGTTATA AATTAGTGAA GAATGATGGC GAATTCCGCT
3351  TACATAACCC AATAAAAGAG CAGGAATTGC GCTCTGATTT AGTAAGAGCA
3401  GAGCAAGCAG AACGAACATT AGAAGCCAAA CAAGTTGAAC AGACTGCTAA
3451  AACACAAACA AGTAAGGCAA GAGTGCGGTC AAGAAGAGCG GTGTTTTCTG
3501  ATCCCCTGCC TGCTCAAAGC CTGTTAAAAG CATTAGAAGC CAAACAAGCT
3551  CTGACTACTG AAACACAAAC AAGTAAGGCA AAAAAAGTGC GGTCAAAAAG
3601  AGCTGCGAGA GAGTTTTCTG ATACCCTGCC TGATCAAATA TTACAAGCCG
3651  CACTTGAGGT TATTGATGCC CAACAGCAAG TGAAAAAAGA ACCTCAAACT
3701  CAAGAGGAAG AAGAGAAAAG ACAACGCAAA CAAAAAGAAT TGATCAGCCG
3751  TTACTCAAAT AGTGCGTTAT CGGAGTTGTC TGCGACAGTA AATAGTATGC
3801  TTTCCGTTCA AGATGAATTG GATCGTCTTT TTGTAGATCA AGCACAATCT
3851  GCCGTGTGGA CAAATATCGC ACAGGATAAA AGACGCTATG ATTCTGATGC
3901  GTTCCGTGCT TATCAGCAGA AAACGAACTT GCGTCAAATT GGGGTGCAAA
3951  AAGCCTTAGA TAATGGACGA ATTGGGGCGG TTTTCTCGCA TAGCCGTTCA
4001  GATAATACCT TTGACGAACA GGTTAAAAAT CACGCGACAT TAGCGATGAT
4051  GTCGGGTTTT GCCCAATATC AATGGGGCGA TTTACAATTT GGTGTAAACG
4101  TGGGTGCGGG AATTAGTGCG AGTAAAATGG CTGAAGAACA AAGCCGAAAA
4151  ATTCATCGAA AAGCGATAAA TTATGGTGTG AATGCAAGTT ATCAGTTCCG
4201  TTTAGGGCAA TTGGGTATTC AGCCTTATTT GGGTGTTAAT CGATATTTTA
```

Fig. 18B

```
4251  TTGAACGTGA AAATTATCAA TCTGAAGAAG TGAAAGTGCA AACACCGAGC
4301  CTTGTATTTA ATCGCTATAA TGCTGGCATT CGAGTTGATT ATACATTTAC
4351  CCCGACAGAT AATATCAGCA TTAAGCCTTA TTTCTTCGTC AATTATGTTG
4401  ATGTTTCAAA CGCTAACGTA CAAACCACTG TAAATCGCAC GATGTTGCAA
4451  CAATCATTTG GGCGTTATTG GCAAAAAGAA GTGGGATTAA AGGCAGAAAT
4501  TTTACATTTC CAACTTTCCG CTTTTATCTC AAAATCTCAA GGTTCACAAC
4551  TCGGCAAACA GCAAAATGTG GGCGTGAAAT TGGGGTATCG TTGGTAAAAA
4601  TCAAC
```

Fig. 18C

Amino acid sequence for NTHi strain TN106 Hap protein (first amino acid to last amino acid):

```
   1  MKKTVFRLNF LTACISLGIV SQAWAGHTYF GIDYQYYRDF AENKGKFTVG
  51  AQDIDIYNKK GEMIGTMMKG VPMPDLSSMV RGGYSTLISE QHLISVAHNV
 101  GYDVVDFGME GENPDQHRFK YKVVKRYNYK SGDRQYNDYQ HPRLEKFVTE
 151  TAPIEMVSYM DGNHYKNFNQ YPLRVRVGSG HQWWKDDNNK TIGDLAYGGS
 201  WLIGGNTFED GPAGNGTLEL NGRVQNPNKY GPLPTAGSFG DSGSPMFIYD
 251  KEVKKWLLNG VLREGNPYAA VGNSYQITRK DYFQGILNQD ITANFWDTNA
 301  EYRFNIGSDH NGRVATIKST LPKKAIQPER IVGLYDNSQL HDARDKNGDE
 351  SPSYKGPNPW SPALHHGKSI YFGDQGTGTL TIENNINQGA GGLYFEGNFV
 401  VKGNQNNITW QGAGVSVGEE STVEWQVHNP EGDRLSKIGL GTLLVNGKGK
 451  NLGSLSVGNG LVVLDQQADE SGQKQAFKEV GIVSGRATVQ LNSADQVDPN
 501  NIYFGFRGGR LDLNGHSLTF ERIQNTDEGA MIVNHNASQT ANITITGNAT
 551  INSDSKQLTN KKDIAFNGWF GEQDKAKTNG RLNVNYQPVN AENHLLLSGG
 601  TNLNGNITQN GGTLVFSGRP TPHAYNHLRR DLSNMEGIPQ GEIVWDHDWI
 651  NRTFKAENFQ IKGGSAVVSR NVSSIEGNWT VSNNANATFG VVPNQQNTIC
 701  TRSDWTGLTT CKTVDLTDKK VINSIPTTQI NGSINLTDNA TVNIHGLAKL
 751  NGNVTLIDHS QFTLSNNATQ TGNIKLSNHA NATVDNANLN GNVNLMDSAQ
 801  FSLKNSHFSH QIQGGEDTTV MLENATWTMP SDTTLQNLTL NNSTVTLNSA
 851  YSAISNNAPR RRRRSLETET TPTSAEHRFN TLTVNGKLSG QGTFQFTSSL
 901  FGYKSDKLKL SNDAEGDYTL SVRNTGKEPV TFGQLTLVES KDNKPLSDKL
 951  TFTLENDHVD AGALRYKLVK NDGEFRLHNP IKEQELRSDL VRAEQAERTL
1001   EAKQVEQTAK TQTSKARVRS RRAVFSDPLP AQSLLKALEA KQALTTETQT
1051   SKAKKVRSKR AAREFSDTLP DQILQAALEV IDAQQQVKKE PQTQEEEEKR
1101   QRKQKELISR YSNSALSELS ATVNSMLSVQ DELDRLFVDQ AQSAVWTNIA
1151   QDKRRYDSDA FRAYQQKTNL RQIGVQKALD NGRIGAVFSH SRSDNTFDEQ
1201   VKNHATLAMM SGFAQYQWGD LQFGVNVGAG ISASKMAEEQ SRKIHRKAIN
1251   YGVNASYQFR LGQLGIQPYL GVNRYFIERE NYQSEEVKVQ TPSLVFNRYN
1301   AGIRVDYTFT PTDNISIKPY FFVNYVDVSN ANVQTTVNRT MLQQSFGRYW
1351  QKEVGLKAEI LHFQLSAFIS KSQGSQLGKQ QNVGVKLGYR W
```

Fig. 19

Nucleotide sequence for NTHi strain 860295 *hap* gene (start codon begins at position 430, stop codon begins at position 4738):

```
   1  GGAGGCAGTG GTGGCGGACA AATTATTGCG ACGGGTACGC CAGAACAAGT
  51  TGCCAAAGTA GAAAGTTCCC ACACCGCCCG CTTCCTTAAA CCGATTTTAG
 101  AAAAACCTTA GAAAAAATGA CCGCACTTTC AGAGAAAACT CACATAAAGT
 151  GCGGTTATTT TATTAGTGAT ATTGTTTTAA TTTTAGTTAT CTGTATAAAT
 201  TACATATAAT ATTAATCCAT CGCAAGATAA GATTACCCAC TAAGTATTAA
 251  GCAAAAACCT AGAAATTTTG GCTTAATTAC TATATAGTTT TACTGCTTTA
 301  TTTTCTTTTG TGCCTTTTAG TTCGTTTTTT TAGCTGAAAT CCCTTAGAAA
 351  ATCACCGCAC TTTTATTGTT CAATAGTCGT TTAACCACGT ATTTTTTAAT
 401  ACGAAAAATT ACTTAATTAA ATAAACATTA TGAAAAAAAC TGTATTTCGT
 451  CTGAACTTTT TAACCGCTTG CATTTCATTA GGGATAGTAT CGCAAGCGTG
 501  GGCAGGTCAC ACTTATTTTG GGATTGACTA CCAATATTAT CGTGATTTTG
 551  CTGAGAATAA AGGGAAGTTT TCAGTTGGGG CTAAAAATAT TGAGGTTTAT
 601  AACAAAGAGG GGACTTTAGT TGGCACATCA ATGACAAAAG CCCCGATGAT
 651  TGATTTTTCT GTGGTGTCGC GAAATGGGGT GGCGGCATTA GTAGGCGATC
 701  AGTATATTGT GAGTGTGGCA CATAACGGTG GATATAATAG CGTTGATTTT
 751  GGAGCAGAAG GTCCAAATCC CGATCAGCAT CGTTTTACTT ATCAAATTGT
 801  AAAAAGAAAT AATTATAAGC CAGGCAAAGA TAACCCTTAT CATGGTGACT
 851  ATCACATGCC TCGTTTGCAC AAATTTGTCA CTGACGCTGA ACCAGCAAAG
 901  ATGACAGACA ATATGAATGG AAAGAACTAC GCTGATTTAA GTAAATATCC
 951  TGATCGTGTG CGTATTGGTA CAGGTGAACA ATGGTGGAGG ACTGATGAAG
1001  AACAAAAGCA AGGAAGTAAG AGTTCATGGC TTGCTGATGC TTATCTGTGG
1051  AGAATAGCAG GTAACACACA TTCACAAAGT GGAGCGGGCA ACGGCACGGT
1101  AAACTTAAGT GGAGATATCA CAAAACCAAA TAACTATGGA CCTCTTCCTA
1151  CGGGTGTTTC GTTTGGAGAT AGTGGTTCTC CAATGTTTAT TTATGATGCA
1201  ATAAAACAAA AATGGCTTAT TAATGGCGTA TTGCAAACTG GTAACCCTTT
1251  CTCGGGAGCT GGAAATGGAT TCCAATTAAT TAGAAAAAAT TGGTTTTATG
1301  ATAATGTCTT TGTAGAAGAT TTGCCTATAA CATTTTTAGA GCCAAGAAGT
1351  AACGGTCATT ATTCATTTAC TTCAAATAAT AATGGAACTG GTACGGTTAC
1401  TCAAACGAAT GAAAAAGTGA GTATGCCTCA ATTTAAAGTC AGAACGGTTC
1451  AGTTATTTAA TGAAGCATTA AAAGAAAAAG ATAAAGAACC TGTTTATGCT
1501  GCAGGTGGTG TAAATGCTTA TAAACCAAGA CTAAATAATG GTAAAAATAT
1551  TTACTTTGGC GATCGAGGAA CAGGAACTTT AACAATTGAA AATAATATAA
1601  ATCAAGGTGC TGGTGGTTTG TATTTTGAGG GTAACTTTAC GGTATCTTCA
1651  GAAAATAATG CAACTTGGCA AGGTGCTGGA GTGCATGTAG GTGAAGACAG
1701  TACTGTTACT TGGAAAGTAA ACGGCGTGGA ACATGATCGC CTTTCTAAAA
1751  TTGGTAAAGG AACGTTGCAT ATTCAAGCAA AAGGTGAAAA CTTAGGCTCA
1801  ATTAGCGTAG GTGACGGCAA AGTCATTTTA GATCAACAAG CCGATGAGAA
1851  CAACCAAAAA CAAGCCTTTA AAGAAGTTGG CATTGTAAGT GGTAGAGCTA
1901  CCGTTCAACT AAATAGTGCA GATCAAGTTG ATCCTAACAA TATTTATTTC
1951  GGATTTCGTG GTGGTCGCTT AGATCTTAAC GGACATTCAT TAACCTTTAA
2001  ACGTATCCAA AATACGGACG AGGGCGCGAT GATTGTGAAC CATAATACAA
```

Fig. 20A

2051 CTCAAGTCGC TAATATTACT ATTACTGGGA ACGAAAGTAT TACTGCTCCA
2101 TCTAATAAAA ATAATATTAA TAAACTTGAT TACAGCAAAG AAATTGCTTA
2151 CAACGGTTGG TTTGGCGAAA CAGATGAAAA TAAACACAAT GGAAGATTAA
2201 ACCTTATTTA TAAACCAACC ACAGAAGATC GTACTTTGCT ACTTTCAGGT
2251 GGAACAAATT TAAAAGGCAA TATTACTCAG GAAGGCGGCA CTTTAGTGTT
2301 TAGTGGTCGC CCAACTCCAC ACGCTTACAA TCATTTAAAT CGCCCAAACG
2351 AGCTTGGGCG ACCTCAAGGC GAAGTGGTTA TTGATGACGA TTGGATCACC
2401 CGCACATTTA AAGCTGAAAA CTTCCAAATT AAAGGCGGAA GTGCGGTGGT
2451 TTCTCGCAAT GTTTCTTCAA TTGAGGGAAA TTGGACAGTC AGCAATAATG
2501 CAAATGCCGC ATTTGGTGTT GTGCCAAATC AGCAAAATAC CATTTGCACG
2551 CGTTCAGATT GGACAGGATT AACGACTTGT AAAACTGTGG ATTTAACCGA
2601 TACAAAAGTT ATTAATTCCA TACCGACAAC ACAAATTAAT GGCTCTATTA
2651 ATTTAACTGA TAATGCAACA GTGAATATTC ATGGTTTAGC AAAACTTAAT
2701 GGTAATGTCA CTTTAATAAA TCATAGCCAA TTTACATTGA GCAACAATGC
2751 CACCCAAACA GGCAATATCC AACTTTCAAA TCACGCAAAT GCAACGGTGG
2801 ACAATGCAAA TTTGAACGGT AATGTGCATT TAACGGATTC TGCTCAATTT
2851 TCTTTAAAAA ACAGCCATTT TTCGCACCAA ATTCAGGGCG ACAAAGACAC
2901 AACAGTGACG TTGGAAAATG CGACTTGGAC AATGCCTAGC GATGCCACAT
2951 TGCAGAATTT AACGCTAAAT AATAGTACTG TTACGTTAAA TTCAGCTTAT
3001 TCAGCTAGCT CAAATAATGC GCCACGTCAC CGCCGTTCAT TAGAGACGGA
3051 AACAACGCCA ACATCGGCAG AACATCGTTT CAACACATTG ACAGTAAATG
3101 GTAAATTGAG CGGGCAAGGC ACATTCCAAT TTACTTCATC TTTATTTGGC
3151 TATAAAAGCG ATAAATTAAA ATTATCCAAT GACGCTGAGG GCGATTACAC
3201 ATTATCTGTT CGCAACACAG GCAAAGAACC CGAAGCCCTT GAGCAATTAA
3251 CTTTGGTTGA AAGCAAAGAT AATAAACCGT TATCAGACAA ACTCAAATTT
3301 ACTTTAGAAA ATGACCACGT TGATGCAGGT GCATTACGTT ATAAATTAGT
3351 GAAGAATAAT GGCGAATTCC GCTTGCATAA CCCAATAAAA GAGCAGGAAT
3401 TGCGCAATGA TTTAGTAAGA GCAGAGCAAG CAGAACGAAC ATTAGAAGCC
3451 AAACAAGTTG AACAGACTGC TGAAACACAA ACAAGTAATG CAAGAGTGCG
3501 GTCAAAAAGA GCGGTGTTTT CTGATACCCT GCCTGATCAA AGCCAGTTAG
3551 ACGTATTACA AGCCGAACAA GTTGAACCGA CTGCTGAAAA ACAAAAAAAT
3601 AAGGCAAAAA AAGTGCGGTC AAAAAGAGCG GTGTTTTCTG ATACCCTGCC
3651 TGATCAAAGC CAGTTAGACG TATTACAAGC CGAACAAGTT GAACCGACTG
3701 CTGAAAAACA AAAAAATAAG GCAAAAAAAG TGCGGTCAAA AAGAGCCGCG
3751 AGAGAGTTTT CTGATACCCC GCTTGATCTA AGCCGGTTAA AGGTATTAGA
3801 AGTCAAACTT GAGGTTATTA ATGCCCAACA GCAAGTGAAA AAAGAACCTC
3851 AAGATCAAGA GAAACAACGC AAACAAAAAG ACTTGATCAG CCGTTATTCA
3901 AATAGTGCGT TATCAGAATT ATCTGCAACA GTAAATAGTA TGCTTTCTGT
3951 TCAAGATGAA TTAGATCGTC TTTTTGTAGA TCAAGCACAA TCTGCCGTGT
4001 GGACAAATAT CGCACAGGAT AAAAGACGCT ATGATTCTGA TGCGTTCCGT
4051 GCTTATCAGC AGAAAACGAA CTTACGTCAA ATTGGGGTGC AAAAAGCCTT
4101 AGCTAATGGA CGAATTGGGG CAGTTTTCTC GCATAGCCGT TCAGATAATA
4151 CTTTTGATGA ACAGGTTAAA AATCACGCGA CATTAACGAT GATGTCGGGT
4201 TTTGCCCAAT ATCAATGGGG CGATTTACAA TTTGGTGTAA ACGTGGGAAC

Fig. 20B

```
4251  GGGAATCAGT GCGAGTAAAA TGGCTGAAGA ACAAAGCCGA AAAATTCATC
4301  GAAAAGCGAT AAATTATGGC GTGAATGCAA GTTATCAGTT CCGTTTAGGG
4351  CAATTGGGCA TTCAGCCTTA TTTTGGAGTT AATCGCTATT TTATTGAACG
4401  TGAAAATTAT CAATCTGAGG AAGTGAAAGT GAAAACGCCT AGCCTTGCAT
4451  TTAATCGCTA TAATGCTGGC ATTCGAGTTG ATTATACATT TACTCCGACA
4501  GATAATATCA GCGTTAAGCC TTATTTCTTC GTCAATTATG TTGATGTTTC
4551  AAACGCTAAC GTACAAACCA CGGTAAATAG CACGGTGTTG CAACAACCAT
4601  TTGGACGTTA TTGGCAAAAA GAAGTGGGAT TAAAAGCGGA AATTTTACAT
4651  TTCCAACTTT CTGCTTTTAT TTCTAAATCT CAAGGTTCGC AACTCGGCAA
4701  ACAGCAAAAT GTGGGCGTGA AATTGGGGTA TCGTTGGTAA AAATCAACAT
4751  AATTGTATCG TTTATTGATA AACAAGGTGG GGCAGATCCC ACCTTTTTTA
4801  TTTCAATAAT GGAACTTTAT TTAATTAAGA GCATCTAAGT AGCACCCCAT
4851  ATAGGGGATT AATTAAGAGG ATTTAATAAT GAATTTAACT AAACTTTTAC
4901  CAGCATTTGC TGCTGCAGTC GTATTATCTG CTTGTGCAAA GGATGCACCT
4951  GAAATGACAA AATCATCTGC GCAAATAGCT GAAATGCAAA CACTTCCAAC
5001  AATCACTGAT AAAACAGTTG TATATTCCTG CAATAAACAA ACTGTAACTG
5051  CCGTGTATCA ATTTGAAAAC CAAGAACCAG TTGCTGCAAT GGTAAGTGTG
5101  GGCGATGGCA TTATTGCGAA AGATTTTACT CGTGATAAAT CACAAAATGA
5151  CTTTACAAGT TTCGTTTCTG GGGATTATGT TTGGAATGTA GATAGTGGCT
5201  TAACGTTAGA TAAATTTGAT TCTGTTGTGC CTGTCAATTT AATTC
```

Fig. 20C

Amino acid sequence for NTHi strain 860295 Hap protein (first amino acid to last amino acid):

```
   1  MKKTVFRLNF LTACISLGIV SQAWAGHTYF GIDYQYYRDF AENKGKFSVG
  51  AKNIEVYNKE GTLVGTSMTK APMIDFSVVS RNGVAALVGD QYIVSVAHNG
 101  GYNSVDFGAE GPNPDQHRFT YQIVKRNNYK PGKDNPYHGD YHMPRLHKFV
 151  TDAEPAKMTD NMNGKNYADL SKYPDRVRIG TGEQWWRTDE EQKQGSKSSW
 201  LADAYLWRIA GNTHSQSGAG NGTVNLSGDI TKPNNYGPLP TGVSFGDSGS
 251  PMFIYDAIKQ KWLINGVLQT GNPFSGAGNG FQLIRKNWFY DNVFVEDLPI
 301  TFLEPRSNGH YSFTSNNNGT GTVTQTNEKV SMPQFKVRTV QLFNEALKEK
 351  DKEPVYAAGG VNAYKPRLNN GKNIYFGDRG TGTLTIENNI NQGAGGLYFE
 401  GNFTVSSENN ATWQGAGVHV GEDSTVTWKV NGVEHDRLSK IGKGTLHIQA
 451  KGENLGSISV GDGKVILDQQ ADENNQKQAF KEVGIVSGRA TVQLNSADQV
 501  DPNNIYFGFR GGRLDLNGHS LTFKRIQNTD EGAMIVNHNT TQVANITITG
 551  NESITAPSNK NNINKLDYSK EIAYNGWFGE TDENKHNGRL NLIYKPTTED
 601  RTLLLSGGTN LKGNITQEGG TLVFSGRPTP HAYNHLNRPN ELGRPQGEVV
 651  IDDDWITRTF KAENFQIKGG SAVVSRNVSS IEGNWTVSNN ANAAFGVVPN
 701  QQNTICTRSD WTGLTTCKTV DLTDTKVINS IPTTQINGSI NLTDNATVNI
 751  HGLAKLNGNV TLINHSQFTL SNNATQTGNI QLSNHANATV DNANLNGNVH
 801  LTDSAQFSLK NSHFSHQIQG DKDTTVTLEN ATWTMPSDAT LQNLTLNNST
 851  VTLNSAYSAS SNNAPRHRRS LETETTPTSA EHRFNTLTVN GKLSGQGTFQ
 901  FTSSLFGYKS DKLKLSNDAE GDYTLSVRNT GKEPEALEQL TLVESKDNKP
 951  LSDKLKFTLE NDHVDAGALR YKLVKNNGEF RLHNPIKEQE LRNDLVRAEQ
1001   AERTLEAKQV EQTAETQTSN ARVRSKRAVF SDTLPDQSQL DVLQAEQVEP
1051   TAEKQKNKAK KVRSKRAVFS DTLPDQSQLD VLQAEQVEPT AEKQKNKAKK
1101   VRSKRAAREF SDTPLDLSRL KVLEVKLEVI NAQQQVKKEP QDQEKQRKQK
1151   DLISRYSNSA LSELSATVNS MLSVQDELDR LFVDQAQSAV WTNIAQDKRR
1201   YDSDAFRAYQ QKTNLRQIGV QKALANGRIG AVFSHSRSDN TFDEQVKNHA
1251   TLTMMSGFAQ YQWGDLQFGV NVGTGISASK MAEEQSRKIH RKAINYGVNA
1301   SYQFRLGQLG IQPYFGVNRY FIERENYQSE EVKVKTPSLA FNRYNAGIRV
1351   DYTFTPTDNI SVKPYFFVNY VDVSNANVQT TVNSTVLQQP FGRYWQKEVG
1401  LKAEILHFQL SAFISKSQGS QLGKQQNVGV KLGYRW
```

Fig. 21

Nucleotide sequence for NTHi strain 3219B *hap* gene (start codon begins at position 388, stop codon begins at position 4561):

```
   1  CCTGAAGACG TTGCTCAAGT TAAAGGCTCT CACACAGCCC GATTCCTTAA
  51  ACCGATTTTA GAAAAACCTT AGAAAAAATG ACCGCACTTT CAGAGAAAAC
 101  TCACATAAAG TGCGGTTATT TTATTAGTGA TATTGTTTTA ATTATTTGTA
 151  TAAATTACAT ACAATATTAA TCCATCGAAA AATAAGATTA CCCACTAAGT
 201  ATTAAGCCAA AACCTAGAAA TTTTGGCTTA ATTACTATAT AATTTTACTC
 251  CTTTATTTTC TTTTGTGCCT TTTAGTTAGT TCGTTTTTTA GCTGAAATCC
 301  CTCAGAAAAT CACCGCACTT TTATTGTTCA ATAGTCGTTT AACCACGTAT
 351  TTTTTAATAC GAAAAATTAC TTAATTAAAT AAACATTATG AAAAAAACTG
 401  TATTTCGTCT TAATTTTCTA ACCGCTTGTA TTTCATTAGG GATAGTATCG
 451  CAAGCGTGGG CAGGTCACAC TTATTTTGGG ATTGACTACC AATATTATCG
 501  TGATTTTGCC GAGAATAAAG GGAAGTTTAC AGTTGGGGCT CAAGATATTG
 551  ATATCTACAA TAAAAAAGGG GAAATGATAG GTACGATGAT GAAAGGTGTG
 601  CCTATGCCTG ATTTATCTTC CATGGTTCGT GGTGGTTATT CAACATTGAT
 651  AAGTGAGCAG CATTTAATTA GCGTCGCACA TAATGTAGGG TATGATGTCG
 701  TTGATTTTGG TATGGAGGGG GAAAATCCAG ACCAACATCG TTTTAAGTAT
 751  AAAGTTGTTA AACGATATAA TTATAAGAGC GGTGATAGAC AATATAATGA
 801  TTATCAACAT CCAAGATTAG AGAAATTTGT AACGGAAACT GCACCTATTG
 851  AAATGGTTTC ATATATGGAT GGTAATCATT ACAAAAATTT TAATCAATAT
 901  CCTTTGCGAG TTAGAGTTGG AAGTGGGCAT CAATGGTGGA AAGACGATAA
 951  TAATAAAACC ATTGGAGACT TAGCCTATGG AGGTTCATGG TTAATAGGTG
1001  GAAATACCTT TGAAGATGGA CCAGCTGGTA ACGGTACATT AGAATTAAAT
1051  GGGCGAGTAC AAAATCCTAA TAAATATGGT CCACTACCTA CGGCAGGTTC
1101  ATTCGGGGAT AGTGGTTCTC CAATGTTTAT TTATGATAAG GAAGTTAAGA
1151  AATGGTTATT AAATGGCGTG TTACGTGAAG GAAATCCTTA TGCTGCAGTA
1201  GGAAACAGCT ATCAAATTAC ACGAAAAGAT TATTTTCAAG GTATTCTTAA
1251  TCAAGACATT ACAGCTAATT TTTGGGATAC TAATGCTGAA TATAGATTTA
1301  ATATAGGGAG TGACCACAAT GGAAGAGTGG CAACAATCAA AAGTACATTA
1351  CCTAAAAAAG CTATTCAGCC TGAACGAATA GTGGGTCTTT ATGATAATAG
1401  CCAACTTCAT GATGCTAGAG ATAAAAATGG CGATGAATCT CCCTCTTATA
1451  AAGGTCCTAA TCCATGGTCG CCAGCATTAC ATCATGGGAA AAGTATTTAC
1501  TTTGGCGATC AAGGAACAGG AACTTTAACA ATTGAAAATA ATATAAATCA
1551  AGGTGCAGGT GGATTGTATT TTGAAGGTAA TTTTGTTGTA AAAGGCAATC
1601  AAAATAATAT AACTTGGCAA GGTGCAGGCG TTTCTGTTGG AGAAGAAAGT
1651  ACTGTTGAAT GGCAGGTGCA TAATCCAGAA GGCGATCGCT TATCCAAAAT
1701  TGGGCTGGGA ACCTTACTTG TTAATGGTAA AGGGAAAAAC TTAGGAAGCC
1751  TGAGTGTCGG TAACGGTTTG GTTGTGTTAG ATCAACAAGC AGATGAATCA
1801  GGTCAAAAAC AAGCCTTTAA AGAAGTTGGC ATTGTAAGTG GTAGAGCTAC
1851  CGTTCAACTA AATAGTGCAG ATCAAGTTGA TCCTAACAAT ATTTATTTCG
1901  GCTTTCGTGG TGGTCGCTTA GATCTTAATG GGCATTCATT AACCTTTGAA
1951  CGTATCCAAA ATACGGATGA AGGCGCGATG ATTGTGAACC ACAACGCTTC
2001  TCAAACCGCA AATATTACGA TTACAGGCAA CGCAACTATT AATTCAGATA
```

Fig. 22A

2051 GCAAACAACT TACTAATAAA AAAGATATTG CATTTAACGG CTGGTTTGGT
2101 GAGCAAGATA AAGCTAAAAC AAATGGTCGT TTAAATGTGA ATTATCAACC
2151 AGTTAATGCA GAAAATCATT TGTTGCTTTC TGGGGGGACA AATTTAAACG
2201 GCAATATCAC GCAAAATGGT GGTACGTTAG TTTTTAGTGG TCGTCCAACG
2251 CCTCATGCTT ACAATCATTT AAGAAGAGAC TTGTCTAACA TGGAAGGTAT
2301 CCCACAAGGC GAAATTGTGT GGGATCACGA TTGGATCAAC CGCACATTTA
2351 AAGCTGAAAA CTTCCAAATT AAAGGCGGAA GTGCGGTGGT TTCTCGCAAT
2401 GTTTCTTCAA TTGAGGGAAA TTGGACAGTC AGCAATAATG CAAATGCCAC
2451 ATTTGGTGTT GTGCCAAATC AGCAAAATAC CATTTGCACG CGTTCAGATT
2501 GGACAGGATT AACGACTTGT AAAACAGTTG ATTTAACCGA TAAAAAAGTT
2551 ATTAATTCCA TACCGACAAC ACAAATTAAT GGTTCTATTA ATTTAACTGA
2601 TAATGCAACA GTGAATATTC ATGGTTTAGC AAAACTTAAT GGTAATGTCA
2651 CTTTAATAGA TCACAGCCAA TTTACATTGA GCAACAATGC CACCCAAGCA
2701 GGCAATATCA AACTTTCAAA TCACGCAAAT GCAACGGTGG ACAATGCAAA
2751 TTTGAACGGT AATGTGAATT TAATGGATTC TGCTCAATTT TCTTTAAAAA
2801 ACAGCCATTT TTCGCACCAA ATCCAAGGTG GGGAAGACAC AACAGTGATG
2851 TTGGAAAATG CGACTTGGAC AATGCCTAGC GATACCACAT TGCAGAATTT
2901 AACGCTAAAT AATAGTACTG TTACGTTAAA TTCAGCTTAT TCAGCTATCT
2951 CAAATAATGC GCCACGCCGT CGCCGCCGTT CATTAGAGAC GGAAACAACG
3001 CCAACATCGG CAGAACATCG TTTCAACACA TTGACAGTAA ATGGTAAATT
3051 GAGCGGGCAA GGCACATTCC AATTTACTTC ATCTTTATTT GGCTATAAAA
3101 GCGATAAATT AAAATTATCC AATGACGCTG AGGGCGATTA CACATTATCT
3151 GTTCGCAACA CAGGCAAAGA ACCCGTGACC TTTGGGCAAT TAACTTTGGT
3201 TGAAAGCAAA GATAATAAAC CGTTATCAGA CAAACTCACA TTCACGTTAG
3251 AAAATGACCA CGTTGATGCA GGTGCATTAC GTTATAAATT AGTGAAGAAT
3301 GATGGCGAAT TCCGCTTACA TAACCCAATA AAAGAGCAGG AATTGCGCTC
3351 TGATTTAGTA AGAGCAGAGC AAGCAGAACG AACATTAGAA GCCAAACAAG
3401 TTGAACAGAC TGCTAAAACA CAAACAAGTA AGGCAAGAGT GCGGTCAAGA
3451 AGAGCGGTGT TTTCTGATCC CCTGCCTGCT CAAAGCCTGT TAAACGCATT
3501 AGAAGCCAAA CAAGCTCTGA CTACTGAAAC ACAAACAAGT AAGGCAAAAA
3551 AAGTGCGGTC AAAAAGAGCT GCGAGAGAGT TTTCTGATAC CCTGCCTGAT
3601 CAAATATTAC AAGCCGCACT TGAGGTTATT GATGCCCAAC AGCAAGTGAA
3651 AAAAGAACCT CAAACTCAAG AGGAAGAAGA GAAAAGACAA CGCAAACAAA
3701 AAGAATTGAT CAGCCGTTAC TCAAATAGTG CGTTATCGGA GTTGTCTGCG
3751 ACAGTAAATA GTATGCTTTC CGTTCAAGAT GAATTGGATC GTCTTTTTGT
3801 AGATCAAGCA CAATCTGCCG TGTGGACAAA TATCGCACAG GATAAAAGAC
3851 GCTATGATTC TGATGCGTTC CGTGCTTATC AGCAGAAAAC GAACTTGCGT
3901 CAAATTGGGG TGCAAAAAGC CTTAGATAAT GGACGAATTG GGGCGGTTTT
3951 CTCGCATAGC CGTTCAGATA ATACCTTTGA CGAACAGGTT AAAAATCACG
4001 CGACATTAGC GATGATGTCT GGTTTTGCCC AATATCAATG GGGCGATTTA
4051 CAATTTGGTG TAAACGTGGG TGCGGGAATT AGTGCGAGTA AAATGGCTGA
4101 AGAACAAAGC CGAAAAATTC ATCGAAAAGC GATAAATTAT GGTGTGAATG
4151 CAAGTTATCA GTTCCGTTTA GGGCAATTGG GTATTCAGCC TTATTTGGGT
4201 GTTAATCGAT ATTTTATTGA ACGTGAAAAT TATCAATCTG AAGAAGTGAA

Fig. 22B

```
4251  AGTGCAAACA CCGAGCCTTG TATTTAATCG CTATAATGCT GGCATTCGAG
4301  TTGATTATAC ATTTACCCCG ACAGATAATA TCAGCATTAA GCCTTATTTC
4351  TTCGTCAATT ATGTTGATGT TTCAAACGCT AACGTACAAA CCACTGTAAA
4401  TCGCACGATG TTGCAACAAT CATTTGGGCG TTATTGGCAA AAAGAAGTGG
4451  GATTAAAGGC AGAAATTTTA CATTTCCAAC TTTCCGCTTT TATCTCAAAA
4501  TCTCAAGGTT CACAACTCGG CAAACAGCAA AATGTGGGCG TGAAATTGGG
4551  GTATCGTTGG TAAAAATCAA CATAATTTTA TCGTTTATTG ATAAACAAGG
4601  TGGGGCAGAT CAAATCCTAC CTTTTTTATT CCAATAATGG AACTTTATTT
4651  TATTAAAGGT ATCTAAGTAG CACCCTATAT AGGGATTAAT TAAGAGGATT
4701  TAATAATGAA TTTAACTAAA ATTTTACCCA CATTTGCTGC TGTAGTCGTA
4751  TTATCTGCTT GTGCAAAGGA TGCACCTGAA ATGACAAAAT CATCTGCGCA
4801 AATAGCTGAA ATGCAAACAC TT
```

Fig. 22C

Amino acid sequence for NTHi strain 3219B Hap protein (first amino acid to last amino acid):

```
   1  MKKTVFRLNF LTACISLGIV SQAWAGHTYF GIDYQYYRDF AENKGKFTVG
  51  AQDIDIYNKK GEMIGTMMKG VPMPDLSSMV RGGYSTLISE QHLISVAHNV
 101  GYDVVDFGME GENPDQHRFK YKVVKRYNYK SGDRQYNDYQ HPRLEKFVTE
 151  TAPIEMVSYM DGNHYKNFNQ YPLRVRVGSG HQWWKDDNNK TIGDLAYGGS
 201  WLIGGNTFED GPAGNGTLEL NGRVQNPNKY GPLPTAGSFG DSGSPMFIYD
 251  KEVKKWLLNG VLREGNPYAA VGNSYQITRK DYFQGILNQD ITANFWDTNA
 301  EYRFNIGSDH NGRVATIKST LPKKAIQPER IVGLYDNSQL HDARDKNGDE
 351  SPSYKGPNPW SPALHHGKSI YFGDQGTGTL TIENNINQGA GGLYFEGNFV
 401  VKGNQNNITW QGAGVSVGEE STVEWQVHNP EGDRLSKIGL GTLLVNGKGK
 451  NLGSLSVGNG LVVLDQQADE SGQKQAFKEV GIVSGRATVQ LNSADQVDPN
 501  NIYFGFRGGR LDLNGHSLTF ERIQNTDEGA MIVNHNASQT ANITITGNAT
 551  INSDSKQLTN KKDIAFNGWF GEQDKAKTNG RLNVNYQPVN AENHLLLSGG
 601  TNLNGNITQN GGTLVFSGRP TPHAYNHLRR DLSNMEGIPQ GEIVWDHDWI
 651  NRTFKAENFQ IKGGSAVVSR NVSSIEGNWT VSNNANATFG VVPNQQNTIC
 701  TRSDWTGLTT CKTVDLTDKK VINSIPTTQI NGSINLTDNA TVNIHGLAKL
 751  NGNVTLIDHS QFTLSNNATQ AGNIKLSNHA NATVDNANLN GNVNLMDSAQ
 801  FSLKNSHFSH QIQGGEDTTV MLENATWTMP SDTTLQNLTL NNSTVTLNSA
 851  YSAISNNAPR RRRRSLETET TPTSAEHRFN TLTVNGKLSG QGTFQFTSSL
 901  FGYKSDKLKL SNDAEGDYTL SVRNTGKEPV TFGQLTLVES KDNKPLSDKL
 951  TFTLENDHVD AGALRYKLVK NDGEFRLHNP IKEQELRSDL VRAEQAERTL
1001   EAKQVEQTAK TQTSKARVRS RRAVFSDPLP AQSLLNALEA KQALTTETQT
1051   SKAKKVRSKR AAREFSDTLP DQILQAALEV IDAQQQVKKE PQTQEEEEKR
1101   QRKQKELISR YSNSALSELS ATVNSMLSVQ DELDRLFVDQ AQSAVWTNIA
1151   QDKRRYDSDA FRAYQQKTNL RQIGVQKALD NGRIGAVFSH SRSDNTFDEQ
1201   VKNHATLAMM SGFAQYQWGD LQFGVNVGAG ISASKMAEEQ SRKIHRKAIN
1251   YGVNASYQFR LGQLGIQPYL GVNRYFIERE NYQSEEVKVQ TPSLVFNRYN
1301   AGIRVDYTFT PTDNISIKPY FFVNYVDVSN ANVQTTVNRT MLQQSFGRYW
1351  QKEVGLKAEI LHFQLSAFIS KSQGSQLGKQ QNVGVKLGYR W
```

Fig. 23

Nucleotide sequence for NTHi strain 1396B *hap* gene (start codon begins at position 313, stop codon begins at position 4546):

```
   1  TGACCGCACT TTCAGAGAAA ACTCACATAA AGTGCGGTTA TTTTATTAGT
  51  GATATTGTTT TAATTTTAGT TATCTGTATA AATTACATAC AATATTAATC
 101  CATCGCAAGA TAAGATTACC CACTAAGTAT TAAGCAAAAA CCTAGAAATT
 151  TTGGCTTAAT TACTATATAG TTTTACTCAT TTATTTTCTT TTGTGCCTTT
 201  TAGTTCGTTT TTTTAGCTGA AATCCCTTAG AAAATCACCG CACTTTTATT
 251  GTTCAATAGT CGTTTAACCA CGTATTTTTT AATACGAAAA ATTACTTAAT
 301  TAAATAAACA TTATGAAAAA AACTGTATTT CGTCTGAATT TTTTAACCGC
 351  TTGCATTTCA TTAGGGATAG TATCGCAAGC GTGGGCAGGT CATACTTATT
 401  TTGGGATTGA CTACCAATAT TATCGTGATT TTGCCGAGAA TAAAGGGAAG
 451  TTCACAGTTG GGGCTAAAAA TATTGAGGTT TACAATAAAA ATGGAAATTT
 501  AGTTGGCACA TCAATGACAA AAGCCCCAAT GATTGATTTT TCCGTGGTGT
 551  CGCGAAATGG GGTGGCGGCA TTGGTGGGCG ATCAGTATAT TGTGAGTGTG
 601  GCACATAATG TAGGCTATAC CAATGTGGAT TTTGGTGCTG AAGGACAAAA
 651  TCCTGATCAA CATCGTTTTA CTTATAAAAT TGTGAAACGG AATAATTATA
 701  AAAACGATCA AACGCATCCT TATGAGAAAG ACTACCACAA CCCACGCTTA
 751  CATAAATTTG TTACGGAAGC CACCCCAATC GATATGACTT CTGATATGAA
 801  CGGCAACAAA TATACAGATA GGACGAAATA TCCCGAACGC GTGCGTATCG
 851  GCTCCGGGTG GCAGTTTTGG CGAAACGATC AAAACAACGG CGACCAAGTT
 901  GCCGGCGCAT ATCATTACCT GACAGCAGGC AATACACACA ACCAAGGCGG
 951  AGCAGGGGGC GGCTGGTCAA GTCTGAGCGG CGATGTGCGC CAAGCGGGCA
1001  ATTACGGCCC CATTCCTATT GCAGGCTCAA GCGGCGACAG CGGTTCGCCT
1051  ATGTTTATTT ATGATGCGGA AAAACAAAAA TGGTTGATTA ACGGCGTATT
1101  GAGGACCGGC AACCCTTGGG CGGGGACAGA GAATACATTC CAACTGGTAC
1151  GCAAGTCTTT TTTTGATGAA ATCCTTGAAA AAGATTTGCG TACATCGTTT
1201  TATAGCCCAT CGGGCAATGG TGCATACACC ATTACAGACA AAGGCGACGG
1251  CAGCGGCATT GTCAAACAAC AAACAGGAAG ACCATCTGAA GTCCGCATCG
1301  GTTTAAAAGA CGACAAATTA CCTGCCGAAG GTAAAGACGA TGTTTACCAA
1351  TACCAAGGTC CAAATATATA CCTGCCTCGT TTGAATAACG GTGGAAACCT
1401  GTATTTCGGA GATCAAAAAA ACGGCACTGT TACCTTATCA ACCAACATCA
1451  ACCAAGGTGC GGGCGGTTTG TATTTTGAGG GTAACTTTAC GGTATCTTCA
1501  GAAAATAATG CAACTTGGCA AGGTGCTGGA GTGCATGTAG GTGAAGACAG
1551  TACTGTTACT TGGAAAGTAA ATGGTGTTGA AAATGATCGC CTTTCTAAAA
1601  TCGGCAAAGG CACATTGCAC GTTAAAGCCA AAGGGGAAAA TAAAGGTTCG
1651  ATCAGCGTAG GCGATGGTAA AGTCATTTTG GAGCAGCAGG CAGACGATCA
1701  AGGCAACAAA CAAGCCTTTA GTGAAATTGG CTTGGTTAGT GGCAGAGGTA
1751  CGGTTCAGTT AAACGATGAC AAGCAATTTA ATACTGATAA ATTTTATTTC
1801  GGCTTCCGTG GTGGTCGCTT AGATCTTAAT GGGCATTCAT TAACCTTTAA
1851  ACGTATCCAA AATACGGATG AGGGAGCAAC GATTGTTAAT CACAATGCCA
1901  CAACAGAATC TACAGTGACC ATTACTGGCA GCGATACCAT TAATGACAAC
1951  ACTGGCGATT TAACCAATAA ACGTGATATT GCTTTTAATG GTTGGTTTGG
2001  TGATAAAGAT GATACTAAAA ATACTGGACG TTTGAATGTT ACTTACAATC
```

Fig. 24A

2051 CGCTTAACAA AGATAATCAC TTCCTTCTAT CAGGTGGAAC AAATTTAAAA
2101 GGCAATATTA CTCAAGACGG TGGCACTTTA GTGTTTAGTG GTCGCCCAAC
2151 ACCACACGCA TACAATCATT TAAATCGCCT AAACGAGCTT GGGCGACCTA
2201 AGGGCGAAGT GGTTATTGAT GACGATTGGA TCAACCGTAC ATTTAAAGCT
2251 GAAAACTTCC AAATTAAAGG CGGAAGTACG GTGGTTTCTC GCAATGTTTC
2301 TTCAATTGAA GGAAATTGGA CAATCAGCAA TAACGCCAAC GCGACATTTG
2351 GTGTTGTGCC AAATCAACAA AATACCATTT GCACGCGTTC AGATTGGACA
2401 GGATTAACGA CTTGTAAAAC AGTTAATTTA ACCGATAAAA AAGTTATTGA
2451 TTCCATACCG ACAACACAAA TTAATGGCTC TATTAATTTA ACTAATAATG
2501 CAACAGTGAA TATTCATGGT TTAGCAAAAC TTAATGGTAA TGTCACTTTA
2551 ATAAATCATA GCCAATTTAC ATTGAGCAAC AATGCCACCC AAACAGGCAA
2601 TATCCAACTT TCAAATCACG CAAATGCAAC GGTGGATAAT GCAAACTTGA
2651 ACGGTAATGT GCATTTAACG GATTCTGCTC AATTTTCTTT AAAAAACAGC
2701 CATTTTTCGC ACCAAATTCA GGGCGACAAA GACACAACAG TGACGTTGGA
2751 AAATGCGACT TGGACAATGC CTAGCGATAC TACATTGCAG AATTTAACGC
2801 TAAATAATAG TACTGTTACG TTAAATTCAG CTTATTCAGC TAGCTCAAAT
2851 AATGCGCCAC GTCACCGCCG TTCATTAGAG ACGGAAACAA CGCCAACATC
2901 GGAAGAACAT CGTTTCAACA CATTGACAGT AAATGGTAAA TTGAGCGGGC
2951 AAGGCACATT CCAATTTACT TCATCTTTAT TTGGCTATAA AAGCGATAAA
3001 ATAAAATTAT CTAATGACGC TGAAGGCGAT TACACATTAG CTGTTCGCGA
3051 CACAGGCAAA GAACCTGTGA CCCTTGAGCA ATTAACTTTA ATTGAAGGCT
3101 TGGATAATCA ACCCTTGCCA GATAAGCTAA AAATTACTTT AAAAAATAAA
3151 CACGTTGATG CGGGTGCATG GCGTTATGAA TTAGTGAAGA AAAACGGCGA
3201 ATTCCGCTTG CATAATCCAA TAAAAGAGCA GGAATTGCGC AATGATTTAG
3251 TAAAAGCAGA GCAAGTAGAA CGAGCATTAG AAGCAAAACA AGCTGAACTG
3301 ACTACTAAAA AACAAAAAAC TGAGGCTAAA GTGCGGTCAA AAAGAGCGGC
3351 GTTTTCTGAT ACCCCGCCTG ATCAAAGCCA GTTAAACGCA TTACAAGCCG
3401 AACTCGAGAC GATTAATGCC CAACAGCAAG TGGCACAAGC GGTGCAAAAT
3451 CAGAAAGTAA CTGCACTTAA CCAAAAGAAC GAGCAAGTTA AAACCACTCA
3501 AGATAAAGCA AATTTAGTCT TGGCAACTGC ATTGGTGGAA AAAGAAACCG
3551 CTCAGATTGA TTTTGCTAAT GCAAAATTAG CTCAGTTGAA TTTAACACAA
3601 CAACTAGAAA AAGCCTTAGC AGTGGCTGAG CAAGCAGAAA AAGAGCGTAA
3651 AGCTCAAGAG CAAGCGAAAA GACAACGCAA ACAAAAAGAC TTGATCAGCC
3701 GTTATTCAAA TAGTGCGTTA TCAGAATTAT CTGCAACAGT AAATAGTATG
3751 CTTTCCGTTC AAGATGAATT AGATCGTCTT TTTGTAGATC AAGCTCAATC
3801 TGCGGTGTGG ACAAATATCT CACAGGATAA AAGACGTTAT GATTCTGATG
3851 CGTTCCGTGC TTATCAGCAG AAAACGAACT TGCGTCAAAT TGGGGTGCAA
3901 AAAGCCTTAG CTAACGGACG AATTGGGGCA GTTTTCTCGC ATAGCCGTTC
3951 AGATAATACT TTTGATGAAC AGGTTAAAAA TCACGCAACA TTAACGATGA
4001 TGTCGGGTTT TGCCCAATAT CAATGGGGTG ATTTACAATT TGGTGTAAAC
4051 GTGGGAACGG GAATTAGTGC GAGTAAAATG GCTGAAGAAC AAAGCCGAAA
4101 AATTCATCGA AAAGCGATAA ATTATGGCGT GAATGCAAGT TATTCGTTCC
4151 ATTTAGGGCA ATTGGGTATT CAGCCTTATT TTGGAGTTAA TCGCTATTTT
4201 ATTGAACGTA AAAATTATCA ATCTGAGGAA GTGAAAGTGC AAACACCGAG

Fig. 24B

```
4251  CCTTGCATTT AATCGCTATA ATGCTGGAGT ACGGGTCGAT TATACGTTTA
4301  CCCCGACAGA GAATATCAGC GTTAAGCCTT ATTTCTTCGT CAATTATGTT
4351  GATGTTTCAA ACGCTAACGT ACAAACCACT GTAAATCGCG CGGTGTTGCA
4401  ACAACCATTT GGACGTTATT GGCAAAAAGA AGTGGGATTA AAAGCGGAAA
4451  TTTTACATTT CCAACTTTCT GCTTTTATTT CTAAATCTCA AGGTTCGCAA
4501  CTCGGTAAAC AGCGAAATAT GGGCGTGAAA TTAGGATATC GTTGGTAAAA
4551  ATCAACATAA TTTTATTCTA ATAATGGAAC TTTATTTAAT TAAAAGTATC
4601  TAAGTAGCAC CCTATAGGGG ATTAATTAAG AGGATTTAAT AATGAATTTA
4651  ACTAAAATTT TACCCGCATT TGCTGCTGCA GTCGTATTAT CTGCTTGTGC
4701  AAAGGATGCA CCTGAAATGA CAAAATCATC TGCGCAAATA GCTGAAATGC
4751  AAACACTTCC AACAATCACT GATAAAACAG TTGTATATTC TTGCAATAAA
4801  CAAACTGTGA CTGCAGTGTA TCAATTTG
```

Fig. 24C

Amino acid sequence for NTHi strain 1396B Hap protein (first amino acid to last amino acid):

```
  1  MKKTVFRLNF LTACISLGIV SQAWAGHTYF GIDYQYYRDF AENKGKFTVG
 51  AKNIEVYNKN GNLVGTSMTK APMIDFSVVS RNGVAALVGD QYIVSVAHNV
101  GYTNVDFGAE GQNPDQHRFT YKIVKRNNYK NDQTHPYEKD YHNPRLHKFV
151  TEATPIDMTS DMNGNKYTDR TKYPERVRIG SGWQFWRNDQ NNGDQVAGAY
201  HYLTAGNTHN QGGAGGGWSS LSGDVRQAGN YGPIPIAGSS GDSGSPMFIY
251  DAEKQKWLIN GVLRTGNPWA GTENTFQLVR KSFFDEILEK DLRTSFYSPS
301  GNGAYTITDK GDGSGIVKQQ TGRPSEVRIG LKDDKLPAEG KDDVYQYQGP
351  NIYLPRLNNG GNLYFGDQKN GTVTLSTNIN QGAGGLYFEG NFTVSSENNA
401  TWQGAGVHVG EDSTVTWKVN GVENDRLSKI GKGTLHVKAK GENKGSISVG
451  DGKVILEQQA DDQGNKQAFS EIGLVSGRGT VQLNDDKQFN TDKFYFGFRG
501  GRLDLNGHSL TFKRIQNTDE GATIVNHNAT TESTVTITGS DTINDNTGDL
551  TNKRDIAFNG WFGDKDDTKN TGRLNVTYNP LNKDNHFLLS GGTNLKGNIT
601  QDGGTLVFSG RPTPHAYNHL NRLNELGRPK GEVVIDDDWI NRTFKAENFQ
651  IKGGSTVVSR NVSSIEGNWT ISNNANATFG VVPNQQNTIC TRSDWTGLTT
701  CKTVNLTDKK VIDSIPTTQI NGSINLTNNA TVNIHGLAKL NGNVTLINHS
751  QFTLSNNATQ TGNIQLSNHA NATVDNANLN GNVHLTDSAQ FSLKNSHFSH
801  QIQGDKDTTV TLENATWTMP SDTTLQNLTL NNSTVTLNSA YSASSNNAPR
851  HRRSLETETT PTSEEHRFNT LTVNGKLSGQ GTFQFTSSLF GYKSDKIKLS
901  NDAEGDYTLA VRDTGKEPVT LEQLTLIEGL DNQPLPDKLK ITLKNKHVDA
951  GAWRYELVKK NGEFRLHNPI KEQELRNDLV KAEQVERALE AKQAELTTKK
1001  QKTEAKVRSK RAAFSDTPPD QSQLNALQAE LETINAQQQV AQAVQNQKVT
1051  ALNQKNEQVK TTQDKANLVL ATALVEKETA QIDFANAKLA QLNLTQQLEK
1101  ALAVAEQAEK ERKAQEQAKR QRKQKDLISR YSNSALSELS ATVNSMLSVQ
1151  DELDRLFVDQ AQSAVWTNIS QDKRRYDSDA FRAYQQKTNL RQIGVQKALA
1201  NGRIGAVFSH SRSDNTFDEQ VKNHATLTMM SGFAQYQWGD LQFGVNVGTG
1251  ISASKMAEEQ SRKIHRKAIN YGVNASYSFH LGQLGIQPYF GVNRYFIERK
1301  NYQSEEVKVQ TPSLAFNRYN AGVRVDYTFT PTENISVKPY FFVNYVDVSN
1351  ANVQTTVNRA VLQQPFGRYW QKEVGLKAEI LHFQLSAFIS KSQGSQLGKQ
1401  RNMGVKLGYR W
```

Fig. 25